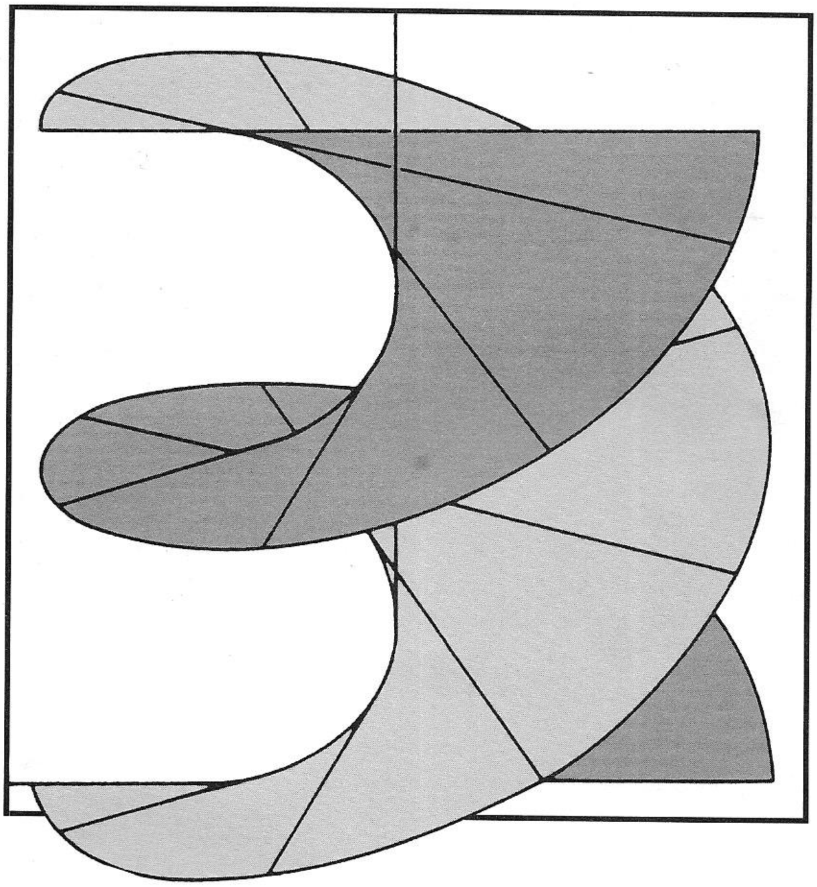
M434 Part I
Mathematics: A Fourth Level Course

M434 DIFFERENTIAL GEOMETRY



PART I
CALCULUS ON EUCLIDEAN SPACE

The Open University

The Open University

Mathematics: A Fourth Level Course

M434 Differential Geometry

Part I Calculus on Euclidean Space

Prepared for the Course Team
by Bob Margolis

Set book

Barrett O'Neill, *Elementary Differential Geometry*, hardback edition (Academic Press, 1966).

It is essential to have this book; the course is based on it and will not make sense without it.

The set book is referred to as *O'Neill*.

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published as *M334* 1976.

First published as *M434* 1993. Reprinted 1996.

Edited, designed and typeset by the Open University using the Open University TEX System.

Printed in the United Kingdom by Hobbs the Printers Ltd, Totton, Hampshire SO40 3WX.

ISBN 0 7492 4770 3

1.2

Contents

Introduction

This text, together with the first two sections of *Part II*, will cover most of the mathematical ideas and methods needed in this course for studying curves and surfaces in three dimensions. Because of the concentration of methods here, unrelieved by any applications, you may find the work rather indigestible at first reading. We hope that a full appreciation of the significance of the ideas will develop as you apply them to curves and surfaces. This introduction is intended to give you some signposts to follow.

The Introduction in *O'Neill* is brief, but important, as it introduces several of the notational conventions used in the course. It also gives a brief reminder of some ideas about functions from $\mathbf{R}$ to $\mathbf{R}$.

It might be as well to say here that in the course as a whole the word 'function', when used without further qualification, should be taken to mean a function from $\mathbf{E}^3$ to $\mathbf{R}$. That is, saying 'f is a function' indicates that

$$f : \mathbf{E}^3 \longrightarrow \mathbf{R}.$$

Before we can hope to describe curves and surfaces, we need a precise description of the 'three-dimensional space' where they will be located. You have already met the vector space $\mathbf{R}^3$ as a mathematical version of the space in which we live—often referred to as (three-dimensional) *Euclidean space*. In Section 1, *O'Neill* refers to this space as $\mathbf{E}^3$. This change of name is not perversity: as will emerge, $\mathbf{E}^3$ is actually rather more than $\mathbf{R}^3$.

As mentioned in *Part 0*, in the historical note, we shall be using sets of axes (Darboux' 'moving frames') attached to each point of the object (curve or surface) that we are studying. Actually, rather than axes, we shall place a set of basis vectors at each point. In order to formalize the notion of 'vector placed at a point' the concept of *tangent vector* is introduced in Section 2. Once we have the idea of a vector placed at a particular point, we shall generalize to placing vectors at each point of $\mathbf{E}^3$ according to some rule. This gives rise to the notion of *vector field*.

These are the 'based vectors' of *MST204*.

Note: As in so much of mathematics, you have to be careful not to read too much into everyday words which are used in a new context. The 'tangent' in 'tangent vector' is a generalization of the usual meaning and does not necessarily imply that the vector is a tangent to a curve or surface in the usual sense of the word. In fact, when dealing with curves we shall have some tangent vectors which are tangents to the curve and some which are not. The new use of 'tangent' will seem strange at first, but you will become used to it with practice.

The key feature of the 'moving frames' approach is the study of the rates of change of the basis vectors as we move around the object being studied. Thus, to make any progress, we shall need careful definitions of what we mean by rates of change, or derivatives, of vectors. We do not attempt to deal with vectors immediately; we start with rates of change of functions. Here we run into the first new idea. The problem is that, in $\mathbf{E}^3$, there are infinitely many ways to set off from a particular point. The rate of change of a function may well depend not only on where you are but also in which direction you set off. Actually, we introduce a further feature: we also permit the rate of change to depend on how fast you set off, as well as in which direction. This is not actually a complication as it permits us to specify the direction and speed of setting off by a vector. The direction is specified by the direction of the vector, and the speed by the size of the vector. Taking all these factors into account leads to the concept of a *directional derivative*. Section 3 defines directional derivatives and proves that they have the linearity and Leibniz (product rule) properties that all our derivatives will have.

Section 4 introduces a formal definition of what is meant in the course by a *curve*. The definition is not quite what you might have expected; curves turn out to be *functions* rather than distorted lines in space.

In Section 5 we return to the theme of directional derivatives of functions. As we indicated above, the derivative of a function can be expected to depend on where we

are (specified by three coordinates) and the vector specifying direction and speed of departure from that point (another three coordinates). Thus the directional derivative will be a function of these six quantities. To give a compact notation for such things, we introduce *1-forms*.

Having introduced 1-forms, Section 6 explores some algebra using 1-forms. Addition and subtraction turn out to be straightforward; multiplication and differentiation are more involved. The ideas from this section will not be applied to anything for some time, so you may well want to skim this section now and come back to it again later.

1-forms will be used in Sections 5–8 of *Part II*.

We are still, at this stage, some way short of the goal of definitions of derivatives of sets of basis vectors. Section 7 makes a step forward by introducing functions

$$\mathbf{E}^n \longrightarrow \mathbf{E}^m,$$

where n and m are usually 2 or 3. We refer to these as *mappings*. You will become familiar with these new objects as you work through the course. Having introduced mappings, we then investigate how they can be combined and differentiated.

Finally, Section 8 provides a very brief summary of this part.

Study advice

One possible breakdown into study weeks is the following.

Week 1 *O'Neill*, Chapter I, Introduction and Sections 1, 2 and 3.

Week 2 *O'Neill*, Chapter I, Sections 4 and 5.

Week 3 *O'Neill*, Chapter I, Sections 6, 7 and 8.

This plan leaves one study week for the first two sections of *Part II* and the associated Tutor-marked Assignment. The suggestion above is exactly that—a suggestion. After working through the first couple of sections you will have a better idea of how fast you can study *O'Neill*. Please allow for the fact that the work becomes less closely connected with previous experience as you go on, and that your progress may become slower as a result.

Do not be tempted to skip exercises. Although there is much powerful theory developed in the course, we expect you to be able to apply the theory to specific examples, and regular practice will help you. We expect that a significant part of your study time will be taken up by the exercises.

The exercises for a particular section may be taken from *O'Neill*, given in this text or both. Solutions for all exercises, regardless of origin, appear at the end of this text. Even if you feel that your solution is completely correct, please read ours; there may well be additional points made in the course of the solution.

Exercises from *O'Neill* are given by page and number.

1 Euclidean space

Read *O'Neill: Introduction and Chapter I, Section 1, pages 1–5.*

Euclidean 3-space You may well have seen this denoted by $\mathbf{R}^3$ rather than $\mathbf{E}^3$. *M203*

Natural coordinate functions This use of x, y and z to represent *functions* may well be new to you. However, the function view does correspond to how these symbols are used in elementary mathematics. For example,

'the value of x at the point $(2,3)$ is 2'

should correspond to the sort of usage of x that is familiar. This phrase is effectively explaining what the value of the function x is at a point in its domain. All that *O'Neill* has done is to formalize this.

Differentiable In differential geometry we shall want our functions to behave 'smoothly'. If you consider a function f from $\mathbf{R}$ to $\mathbf{R}$, it will have a graph without 'sharp points' if we require that its derivative f' exists everywhere. The existence of f' on the domain of f ensures that the graph has a tangent everywhere. If we also require f' to be continuous, we shall ensure that the tangent to the graph cannot make any sudden changes of direction. Going further, requiring f to have derivatives of all possible orders should ensure that the graph of f is 'smooth'. The definition on page 4 generalizes this idea to $\mathbf{E}^3$.

The comments on the definition (page 5) simply point out that we really need to worry about f and its derivatives only on the region of $\mathbf{E}^3$ in which we are interested.

The region will always be an open subset of $\mathbf{E}^3$.

The final comment: that some of the concepts are 'over-elaborate' in one dimension is quite important. Actually, the differences between a number of the concepts apparently vanish in the one-dimensional case. Thus, in some cases, generalizing a familiar idea about functions from $\mathbf{R}$ to $\mathbf{R}$ requires not one but several new definitions.

Exercise 1.1 *O'Neill*, page 5, Exercise 1.

Exercise 1.2 *O'Neill*, page 5, Exercise 2.

Exercise 1.3 *O'Neill*, page 6, Exercise 3.

Exercise 1.4 *O'Neill*, page 6, Exercise 4. Note that this exercise defines the composite of a function

$$h : \mathbf{E}^3 \longrightarrow \mathbf{R}$$

with three functions

$$g_1, g_2, g_3 : \mathbf{E}^3 \longrightarrow \mathbf{R}.$$

[*Solutions on page 27*]

2 Tangent vectors

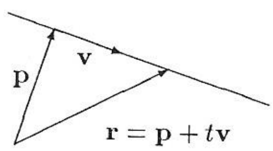
Read *O'Neill: Chapter I, Section 2, pages 6–10.*

Tangent vector The reason for the apparently curious choice of name for this concept will emerge when we study surfaces, where most of the tangent vectors will also be tangents (in the intuitive sense) to surfaces. For now, it is important to realize that a tangent vector may not be (geometrically) a tangent to anything obvious. The important thing is to visualize $\mathbf{v_p}$ as a vector parallel to $\mathbf{v}$, but with its blunt end at the point $\mathbf{p}$ rather than at the origin.

We shall be studying geometric objects from the point of view of someone living on the object, with a personal set of axes. The concept of a tangent vector is exactly what we shall need to define axes at a particular point.

We shall make another use of tangent vectors. A tangent vector $\mathbf{v_p}$ defines a line through $\mathbf{p}$ pointing in the direction of $\mathbf{v}$ by the equation

$$\mathbf{r} = \mathbf{p} + t\mathbf{v}, \quad t \in \mathbf{R}.$$



Tangent space This definition formalizes the idea of taking an egocentric view of the world! One way of looking at it is to consider the 'global' coordinate system as being imposed by an outsider. Then, if you are located at $\mathbf{p}$,

$$T_{\mathbf{p}}(\mathbf{E}^3)$$

represents your personal copy of $\mathbf{E}^3$, with its 'origin' at $\mathbf{p}$.

Vector field This concept will be used to formalize the idea of carrying a set of axes around. Instead of actually moving, say, the x-axis around, we assume that each point $\mathbf{p}$ in space has been provided with the unit vector $(1,0,0)_{\mathbf{p}}$ to indicate the direction of the (local) x-axis. The *function*

$$U_1 : \mathbf{p} \longmapsto (1,0,0)_{\mathbf{p}},$$

which provides the unit vector at each point, is an example of a vector field. Note the difference between U_1, which is a function, and $U_1(\mathbf{p})$, which is a particular tangent vector at a particular point.

Pointwise principle This principle is nothing more than the fact that to define a function you must specify its value at each point in the domain.

Natural frame field This set of vector fields essentially provides each point in $\mathbf{E}^3$ with a copy of the standard basis. Lemma 2.5 says that, because a tangent vector at any point can be expressed in terms of $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ at that point, a corresponding statement is true of a vector field. *M203*

We shall discuss frame fields in general in *Part II*; for now, the only frame field that we shall use is the natural frame field. It is worth emphasizing that 'the natural frame field' means the *three* vector fields U_1,U_2 and U_3.

If you think about how actually to specify a vector field, Lemma 2.5 becomes natural, not to say inevitable. Specifying a vector field involves explaining which tangent vector should be attached to the point

$$\mathbf{p} = (p_1, p_2, p_3),$$

whatever the coordinates of $\mathbf{p}$ happen to be. Thus we would have to explain how to calculate the components of $\mathbf{v}$ from the coordinates of $\mathbf{p}$. That is, we have to specify the functions that *O'Neill* calls v_1, v_2 and v_3.

The identity

$$V = v_1U_1 + v_2U_2 + v_3U_3$$

explains the importance of the natural frame field: it forms a basis for vector fields on $\mathbf{E}^3$.

If you regard a vector field V as a way of persuading $\mathbf{E}^3$ to grow a (straight) hair at each point, then requiring V to be differentiable is to demand that the direction and size of the hairs varies smoothly as you move from point to point of $\mathbf{E}^3$.

Finally, a comment on $\mathbf{R}^3$ versus $\mathbf{E}^3$. We would regard $\mathbf{R}^3$ as a three-dimensional, real vector space and $\mathbf{E}^3$ as that vector space, *together with all its tangent vectors.*

Exercise 2.1 *O'Neill*, page 10, Exercise 1 (part (a) only).

Exercise 2.2 *O'Neill*, page 10, Exercise 2.

Exercise 2.3 *O'Neill*, page 10, Exercise 3. Your answers should be expressed in terms of x, y, z, U_1, U_2 and U_3.

Exercise 2.4 *O'Neill*, page 11, Exercise 5.

Note: This exercise shows that the natural frame field is not the only possible basis for vector fields on $\mathbf{E}^3$. The phrase *Euclidean* coordinate functions was used earlier for the v_i to indicate that they were the coordinate functions corresponding to using the *natural* frame field as a basis.

[*Solutions on page 27*]

3 Directional derivatives

Read *O'Neill: Chapter I, Section 3, pages 11–14.*

This section marks the start of the process of defining rates of change of objects (functions, tangent vectors etc.) as we move around in $\mathbf{E}^3$.

Directional derivatives There is a fundamental problem that arises in generalizing the idea of derivative from one dimension to several (here three) dimensions. When we want to find the rate of change of a function

$$f : \mathbf{R} \longrightarrow \mathbf{R}$$

at a point $a \in \mathbf{R}$, we look at the values of f at points 'near' a. A point in $\mathbf{R}$ near a can be written $a + h$, where h is either positive or negative. The derivative of f at a is found by calculating the limit of

M203

$$\frac{f(a+h) - f(a)}{h}$$

as h tends to zero (provided that the limit exists).

Suppose now that we want the 'derivative' of a function

$$f : \mathbf{E}^3 \longrightarrow \mathbf{R}.$$

A point 'near' $\mathbf{p}$ will be of the form $\mathbf{p} + \mathbf{v}$, where $\mathbf{v}$ is a *vector*. The number h in the one-dimensional case has been replaced by the vector $\mathbf{v}$. Since division by vectors is not defined, we cannot simply translate the limit definition above into the new situation.

We have to accept that the rate of change of f will depend not only on where we are (the point $\mathbf{p}$), but also on the direction in which we move away from $\mathbf{p}$. Having accepted this change from the one-dimensional case, we make a further change: we allow the rate of change of f to depend on how 'fast' we set off from $\mathbf{p}$, as well as on the direction.

Now, the direction and speed with which we set off from $\mathbf{p}$ can conveniently be defined by giving a tangent vector $\mathbf{v_p}$ based at $\mathbf{p}$, since $\mathbf{p}$ and $\mathbf{v}$ define a straight line

$$\mathbf{r} = \mathbf{p} + t\mathbf{v}, \quad t \in \mathbf{R},$$

where $t = 0$ corresponds to the point $\mathbf{p}$.

We can now define a composite function from $\mathbf{R}$ to $\mathbf{R}$ by calculating values of f along this line:

$$t \longmapsto \mathbf{p} + t\mathbf{v} \longmapsto f(\mathbf{p} + t\mathbf{v}).$$

This composite is just a function on $\mathbf{R}$, and so can be differentiated with respect to t in the usual way. The value of this (ordinary) derivative at $t = 0$ is defined to be the derivative of f with respect to the tangent vector $\mathbf{v_p}$. The formal statement of the above discussion is Definition 3.1.

Note: If you prefer to denote differentiation by dashes, you may prefer the following form of the definition:

$$\mathbf{v_p}[f] = (f(\mathbf{p} + t\mathbf{v}))'(0).$$

The dash represents differentiation with respect to t.

There is a strong link between the newly defined directional derivative and the partial derivatives that were discussed in *Part 0*.

Suppose that we look at the tangent vector

$$\mathbf{v_p} = (1, 0, 0)_\mathbf{p}.$$

The definition of $\mathbf{v_p}[f]$ now gives

$$\begin{aligned}\mathbf{v_p}[f] &= (f(\mathbf{p} + t\mathbf{v}))'(0)\\ &= (f(\mathbf{p} + t(1,0,0)))'(0)\\ &= (f(p_1 + t, p_2, p_3))'(0).\end{aligned}$$

In the last line above, only the first coordinate, $p_1 + t$, is varying; the other two are constant. In other words, the last line is

$$\frac{\partial f}{\partial x}$$

evaluated at $t = 0$, in other words at $\mathbf{p}$.

Hence,

$$(1,0,0)_\mathbf{p}[f] = \frac{\partial f}{\partial x}(\mathbf{p}).$$

Similarly,

$$(0,1,0)_\mathbf{p}[f] = \frac{\partial f}{\partial y}(\mathbf{p}),$$

$$(0,0,1)_\mathbf{p}[f] = \frac{\partial f}{\partial z}(\mathbf{p}).$$

Thus, the values of the partial derivatives at a point are the directional derivatives with respect to

$$(1,0,0)_\mathbf{p}, \quad (0,1,0)_\mathbf{p} \quad \text{and} \quad (0,0,1)_\mathbf{p}.$$

This link may well be what is meant by the comment in the paragraph immediately following the definition on page 11 of *O'Neill*.

The worked example following the definition shows how to calculate directional derivatives from first principles. However, Lemma 3.2 provides an alternative method which is of considerable importance.

The proof of Lemma 3.2 uses a generalized form of the chain rule (composite rule) for differentiating composite functions. You have met the rule for functions from $\mathbf{R}$ to $\mathbf{R}$, probably in the form

M101 and *M203*

$$(f \circ g)' = f'(g) \times g'.$$

Here, we have a composite

$$t \longmapsto (p_1 + tv_1, p_2 + tv_2, p_3 + tv_3) \longmapsto f(\mathbf{p} + t\mathbf{v}).$$

If we define

$$\begin{aligned} g_1(t) &= p_1 + tv_1, \\ g_2(t) &= p_2 + tv_2, \\ g_3(t) &= p_3 + tv_3, \end{aligned}$$

then we are trying to differentiate the composite

$$f(g_1, g_2, g_3).$$

The version of the chain rule that applies in this case is

$$(f(g_1, g_2, g_3))' = \frac{\partial f}{\partial x} g_1' + \frac{\partial f}{\partial y} g_2' + \frac{\partial f}{\partial z} g_3'.$$

We shall make quite heavy use of this result. You should note it, and note also its similarity to the 'single variable' case.

It is this result that is applied to prove Lemma 3.2.

Once the definition of directional derivative and a practical method of calculation are established, *O'Neill* goes on to show that directional derivatives have the linearity and Leibnizian (product) properties that you should be familiar with from ordinary derivatives. You might like to compare parts 2 and 3 of Theorem 3.3 (*O'Neill*, page 12) with

$$\begin{aligned} (af + bg)' &= af' + bg', \\ (fg)' &= f'g + fg', \end{aligned}$$

for functions f, g from $\mathbf{R}$ to $\mathbf{R}$ and $a, b \in \mathbf{R}$.

There is no result in 'ordinary' calculus corresponding to Theorem 3.3(1) because ordinary differentiation is always carried out in a standard direction and at standard (unit) speed.

Note that the right-hand side of Theorem 3.3(3) is NOT

$$\mathbf{v_p}[f]\, g + f\, \mathbf{v_p}[g].$$

To see why not, consider what sort of object

$$\mathbf{v_p}[fg]$$

is. The directional derivative of a function with respect to a particular tangent vector is a real number. However,

This type of analysis will often help you to reconstruct such formulas from a general knowledge of the pattern for the Leibniz property.

$$\mathbf{v_p}[f]\, g + f\, \mathbf{v_p}[g]$$

is the sum of two terms, each of which is a function multiplied by a real number. Thus

$$\mathbf{v_p}[f]\, g + f\, \mathbf{v_p}[g]$$

is a function, not a real number. On the other hand, evaluating the functions in each term at $\mathbf{p}$ gives

$$\mathbf{v_p}[f]\, g(\mathbf{p}) + f(\mathbf{p})\, \mathbf{v_p}[g],$$

which is a real number.

The next step that *O'Neill* takes is similar (but not identical) to the process of going from the value of a derivative at a particular point to the idea of a derived function. Thus, in elementary calculus, after finding that the derivative of

Since x is a *function*, this way of defining f is permissible.

$$f = x^2$$

at 2 is 4, at 3 is 6 and so on, we go on to define the derived function

$$f' = 2x,$$

which gives a rule for calculating the derivative everywhere.

Our goal is to give an expression that will enable us to calculate the value of the directional derivative of a function with respect to any tangent vector at any point. In this section we go only part of the way towards this goal; the rest of the work is done in Section 5.

Instead of considering the directional derivative of a function f with respect to a single tangent vector, we now consider the directional derivatives with respect to a whole collection of tangent vectors, one at each point of $\mathbf{E}^3$. The collection of tangent vectors is provided by a vector field V. If you inspect the paragraph beginning half-way down page 13 carefully, you will see that $V[f]$ is defined as a composite function

$$\mathbf{p} \longmapsto V(\mathbf{p}) \longmapsto V(\mathbf{p})[f].$$

Since the domain of this composite is $\mathbf{E}^3$ and the codomain is $\mathbf{R}$, we have that:

the derivative of a function with respect to a vector field is a function.

The example that immediately follows the definition is extremely important. From remarks that we made earlier,

$$(1,0,0)_{\mathbf{p}}[f] = \frac{\partial f}{\partial x}(\mathbf{p}).$$

Since $U_1(\mathbf{p}) = (1,0,0)_{\mathbf{p}}$, we have $U_1[f]$ defined as the composite

$$\mathbf{p} \longmapsto U_1(\mathbf{p}) \longmapsto (1,0,0)_{\mathbf{p}}[f] = \frac{\partial f}{\partial x}(\mathbf{p}).$$

Since the function defined by

$$\mathbf{p} \longmapsto \frac{\partial f}{\partial x}(\mathbf{p})$$

is, by definition, written

$$\frac{\partial f}{\partial x},$$

we have the result

$$U_1[f] = \frac{\partial f}{\partial x}$$

and two similar results for $U_2[f]$ and $U_3[f]$.

These results give a sound basis for thinking of the partial derivatives of a function as being the derivatives in the x, y and z directions.

Now that *O'Neill* has made the definition of $V[f]$, the next step is to show that the newly defined derivative also has the linearity and Leibnizian properties. This is the content of Corollary 3.4. Note that, this time, the Leibniz property (Corollary 3.4(3)) is the exact analogue of

$$(fg)' = f'g + fg'.$$

The results of Corollary 3.4 provide the basis for calculating directional derivatives with respect to vector fields. (Parts 1 and 2 are probably the most used.) The following worked example may help to illustrate the point.

Example Suppose that the function $f : \mathbf{E}^3 \longrightarrow \mathbf{R}$ and the vector field V on $\mathbf{E}^3$ are defined by:

$$f = xe^y, \quad V = z^2U_1 + xzU_2 + x^3U_3.$$

Then

$$\begin{aligned} V[f] &= (z^2U_1 + xzU_2 + x^3U_3)[xe^y] \\ &= z^2U_1[xe^y] + xzU_2[xe^y] + x^3U_3[xe^y] \quad \text{(by Corollary 3.4 (1))} \\ &= z^2\frac{\partial}{\partial x}(xe^y) + xz\frac{\partial}{\partial y}(xe^y) + x^3\frac{\partial}{\partial z}(xe^y) \quad (\text{using } U_1[f] = \partial f/\partial x \text{ etc.}) \\ &= z^2\, e^y + xz\, xe^y + x^3 \times 0 \\ &= ze^y(z + x^2). \end{aligned}$$

The example in *O'Neill* preceding Remark 3.5 uses exactly the same technique: application of linearity followed by

$$U_1[f] = \frac{\partial f}{\partial x}$$

etc. ■

You should note the contents of Remark 3.5, but the course texts will tend to use the full form $\mathbf{v_p}$ rather more than *O'Neill* does. In particular, we shall often retain the full form, $\mathbf{v_p}$, including the point of application $\mathbf{p}$, whereas *O'Neill* tends to retain only the vector part.

Exercise 3.1 *O'Neill*, page 14, Exercise 1.

Exercise 3.2 *O'Neill*, page 15, Exercise 2.

Exercise 3.3 The vector field V and function f are defined by

$$V = 2U_1 - U_2 + 3U_3, \quad f = e^x \cos y.$$

Calculate $V[f]$.

Exercise 3.4 *O'Neill*, page 15, Exercise 3.

Exercise 3.5 Suppose that the vector field V has Euclidean coordinate functions v_1, v_2 and v_3, so that

$$V = v_1U_1 + v_2U_2 + v_3U_3.$$

Find $V[x_i]$ for each coordinate function x_i, and hence show that

$$V = \sum_{i=1}^{3} V[x_i]U_i.$$

Exercise 3.6 If V and W are two vector fields such that

$$V[f] = W[f]$$

for every real-valued function f on $\mathbf{E}^3$, show that $V = W$. (*Hint*: The result of the last exercise may be useful.)

[*Solutions on page 29*]

4 Curves

Read *O'Neill: Chapter I, Section 4, pages 15–21.*

This section formally defines curves, which is the first type of geometric object that you will study in detail in the course; the other being surfaces.

Curve The definition of curve is, perhaps, not quite what you might have expected. It is the function α which is the curve; the object in $\mathbf{E}^3$ that you probably visualize when thinking about a curve is referred to as the *route* of the curve. (The route is the image set of the function α.)

The requirement for the domain of a curve to be an interval, together with requiring a curve to be differentiable, ensure that the route of a curve is connected ('all in one piece').

Much of the language used to discuss curves is based on the idea of a moving point tracing out the route of the curve. The choice of t in

$$\alpha : t \longmapsto \alpha(t) = (\alpha_1(t), \alpha_2(t), \alpha_3(t))$$

suggests 'time' as the variable. We shall talk about 'speed' and 'velocity' in connection with curves.

Example 4.2 Examples 4.2(1) and 4.2(2) in *O'Neill* are particularly important and will reappear from time to time.

Example 4.2(2) The distance of a point in the xy-plane from the origin can be found by evaluating the function

$$\sqrt{x^2 + y^2}$$

at the point in question. Thus

$$(\sqrt{x^2 + y^2})(a\cos t, a\sin t, 0) = \sqrt{a^2\cos^2 t + a^2\sin^2 t} = a$$

so all points on the route of the first curve are on a circle of radius a.

Example 4.2(3) The justification for the description of the route of the curve given in *O'Neill* is contained in Exercise 1 below.

Example 4.2(5) 'Brute-force' plotting of points is only rarely helpful for curves in $\mathbf{E}^3$, partly due to the difficulty of sketching points on such routes on two-dimensional paper.

Velocity of curve This section of *O'Neill* introduces the concept of velocity and gives it a geometric interpretation. The geometric link might be a little clearer if the two diagrams on pages 17 and 18 were interchanged: the second is before the limit is taken, the first is after.

It is important to note that the velocity vector

$$(\alpha_1'(t), \alpha_2'(t), \alpha_3'(t))$$

is attached to the point $\alpha(t)$ on the route; it is a tangent vector in the sense that we have recently defined.

There is also a link with the idea of vector field. The process of finding velocities at all the points on the route of a curve amounts to defining a vector field

$$\alpha_1' U_1 + \alpha_2' U_2 + \alpha_3' U_3,$$

but *only* for points on the route of the curve.

The term 'speed' is mentioned in this section of *O'Neill* without explanation. The definition is straightforward: it is the magnitude of the velocity. Thus, at $\alpha(t)$, the speed is

$$\sqrt{(\alpha_1'(t))^2 + (\alpha_2'(t))^2 + (\alpha_3'(t))^2}.$$

Since $||\mathbf{v}||$ is used in this course to denote the size (magnitude) of the vector $\mathbf{v}$, this result can be written

You may have used $|\mathbf{v}|$ in the past.

$$\text{Speed} = ||\alpha|| = \sqrt{(\alpha_1'(t))^2 + (\alpha_2'(t))^2 + (\alpha_3'(t))^2}.$$

Reparametrization This is an important technical process, although you would be forgiven for not seeing why at this point in the course!

To give an indication of its importance, we must refer back to the method that we said would be used to study curves: each point will be provided with a 'local' basis of unit vectors and the change in this basis from place to place will give information about the shape of the curve.

The velocity vector provides *a* vector at each point on the route of the curve, but it is not necessarily of unit length. If we could arrange that the length of the velocity vector were 1 at each point (that is, the curve had unit speed everywhere), we should have a start for our basis of unit vectors. The later work on reparametrization will show that we can always arrange for the route of a curve to be traversed at unit speed.

This section makes a start by showing what happens to the velocity when the parametrization is changed. *Part II*, Section 2, will complete the process by showing how to change the parametrization to achieve unit speed.

The key step in Lemma 4.5 involves three applications of the usual chain rule formula. If you inspect the result carefully, you will see that it can be written

$$\beta = \alpha \circ h \quad \Longrightarrow \quad \beta' = \alpha'(h) \times h',$$

which is a direct analogue of the familiar form of the chain rule.

Important: Lemma 4.6 is apparently there simply because we have invented a new sort of tangent vector—velocities of curves—and we can now use such tangent vectors to differentiate functions. In a sense this is so. However, Lemma 4.6 will be of *vital* importance when we come to study surfaces, so it is worth taking some time to appreciate exactly what it says.

As *O'Neill* says on page 20, finding $\alpha'(t)[f]$ involves finding the rate of change of f along the straight line through $\alpha(t)$ in the direction of $\alpha'(t)$. The lemma shows that the rate of change of f along the tangent is the same as the rate of change along the curve itself. In one way this is not surprising: the curve and the tangent point in the same direction and the length of the tangent represents the speed of progress along the curve.

Where, as here, it causes no confusion, we say 'along the curve' rather than insist on the more precise 'along the route of the curve'.

Lemma 4.6 is of only theoretical interest at the moment. Later, however, we shall encounter functions which are defined only on a curve or surface and not on the whole of $\mathbf{E}^3$. In such cases we cannot use the definition to find directional derivatives because, given $\mathbf{v_p}$, the function in question may not be defined along the whole of the straight line $\mathbf{p} + t\mathbf{v}$. In such cases we use Lemma 4.6. If we can find a curve α, such that

$$\alpha(0) = \mathbf{p}, \quad \alpha'(0) = \mathbf{v_p},$$

then, for a function f, we have

$$\mathbf{v_p}[f] = (f(\alpha(t)))'(0).$$

Paraphrasing the result of Lemma 4.6: in the definition of directional derivative, the straight line $\mathbf{p} + t\mathbf{v}$ can be replaced by *any* curve α through $\mathbf{p}$ with velocity $\mathbf{v_p}$ at $\mathbf{p}$.

The final remarks in this section of *O'Neill* are not of great importance for this course and you need read them only cursorily.

Exercise 4.1 Let $\alpha : (0, \pi/2) \longrightarrow \mathbf{E}^3$ be the curve described in Example 4.2(3) of *O'Neill*, that is,

$$\alpha(t) = (2\cos^2 t, \sin 2t, 2\sin t), \quad 0 < t < \pi/2.$$

(a) Find the velocity vector of α for arbitrary t and also for $t = \pi/4$.

(b) Justify *O'Neill*'s statements about the curve by doing the following.

(i) Show that $\alpha(t)$ has length 2 for all values of t. (This shows that the route of α lies on the sphere of radius 2, centre the origin.)

(ii) Show that $\alpha(t)$ is always 1 unit from the point $(1, 0, 2\sin t)$. (This shows that the route of α lies on the cylinder of radius 1 whose axis is the line $x = 1, y = 0$.)

(c) Find the coordinate functions of the curve $\beta = \alpha(h)$, where h is the function defined by

$$h(s) = \arcsin s, \quad 0 < s < 1.$$

Exercise 4.2 *O'Neill*, page 21, Exercise 4. (2)

Exercise 4.3 *O'Neill*, page 21, Exercise 5. Note that the phrase 'find a straight line ...' means 'write down the equation in the form $\alpha(t) = \ldots$ for a suitable function α'.

Exercise 4.4 *O'Neill*, page 21, Exercise 7.

Exercise 4.5 *O'Neill*, page 22, Exercise 9.

[*Solutions on page 31*]

5 1-Forms

Read *O'Neill: Chapter I, Section 5, pages 22–25.*

Erratum *O'Neill*, page 25, final line before the exercises:

for $\ldots(p_2^2 + 1)v_3$ read $\ldots(p_2^2 + 2)v_3$.

Do not worry if you have not met the definition of differential given at the start of this section.

1-forms In Section 3 we mentioned that this section would complete the task of giving a formula for calculating the directional derivative of a function at any point with respect to any tangent vector.

Consider, for a moment, what such a formula would have to look like. Given a function f and a tangent vector $\mathbf{v}_\mathbf{p}$, the directional derivative would depend on the three components p_1, p_2 and p_3 of $\mathbf{p}$ *and* on the three components v_1, v_2 and v_3 of $\mathbf{v}$.

For example, let

$$f = xyz \quad \text{and} \quad \mathbf{v_p} = (v_1, v_2, v_3)_{(p_1, p_2, p_3)},$$

then, working from Lemma3.2,

$$\begin{aligned}\mathbf{v_p}[f] &= \frac{\partial f}{\partial x}(\mathbf{p})\, v_1 + \frac{\partial f}{\partial y}(\mathbf{p})\, v_2 + \frac{\partial f}{\partial z}(\mathbf{p})\, v_3 \\ &= (yz)(\mathbf{p})\, v_1 + (xz)(\mathbf{p})\, v_2 + (xy)(\mathbf{p})\, v_3 \\ &= p_2 p_3 v_1 + p_1 p_3 v_2 + p_1 p_2 v_3.\end{aligned}$$

Thus, as expected, $\mathbf{v_p}[f]$ is a 'function of six variables'.

What we are aiming at is a formula (which will be called df) which will produce the correct value of the directional derivative when $\mathbf{v_p}$ is correctly substituted.

We already know how to produce the p_i from a point $\mathbf{p}$. We use the coordinate functions x, y and z. What we now need are functions to play a similar role in extracting the v_i from a tangent vector. That is precisely what is done by

$$dx, \quad dy \quad \text{and} \quad dz.$$

For the function defined above, we have

$$df = (yz)\, dx + (xz)\, dy + (xy)\, dz.$$

The concept of 1-form arises from the need to express directional derivatives concisely. However, they are not *defined* quite in the way that the need arises.

Because (Theorem 3.3(1)) directional derivatives behave linearly with respect to the involvement of the tangent vector, the expression that we arrive at for a general directional derivative must involve the v_i in a linear manner. Also, for a particular tangent vector, the value of the directional derivative is a real number. Thus, once we have defined it properly, df will have to be a function whose domain is the set of all possible tangent vectors, whose codomain is $\mathbf{R}$ and which is linear with respect to tangent vectors at each point.

If you check the example above, you will see that this is the case.

We have two properties of df: it is a function from tangent vectors to reals and, at any point $\mathbf{p}$, it is linear. As is so often the case in mathematics, these two properties are extracted and made the definition of a new concept. Here the concept is that of a 1-form. The formal definition is given in Definition 5.1.

Even though df is not properly defined yet!

Having given a definition of 1-forms (5.1) that does not mention directional derivatives, there is now the task of showing that 1-forms are the right concept for expressing general 'derivatives'. This is the point of Definition 5.2.

We have been discussing a way of representing the function

$$\mathbf{v_p} \longmapsto \mathbf{v_p}[f]$$

as a 'formula'. Definition 5.2 makes a start by giving a name, df, to

$$\mathbf{v_p} \longmapsto \mathbf{v_p}[f].$$

A short argument is needed to show that df is a 1-form. (Actually, the definition of 1-form was carefully framed so as to make df a 1-form.)

The argument is in the paragraph immediately after the definition.

Example 5.3(1) Having introduced a new concept, some simple examples are in order. The simplest non-constant functions on $\mathbf{E}^3$ are probably the coordinate functions. It is worth expanding the working in Example 5.3 a little. Since x is a function, we can calculate

$$\begin{aligned}\mathbf{v_p}[x] &= \frac{\partial x}{\partial x}(\mathbf{p})\, v_1 + \frac{\partial x}{\partial y}(\mathbf{p})\, v_2 + \frac{\partial x}{\partial z}(\mathbf{p})\, v_3 \\ &= 1\, v_1 + 0\, v_2 + 0\, v_3 \\ &= v_1.\end{aligned}$$

Hence

$$dx : \mathbf{v_p} \longmapsto v_1.$$

Similarly,

$$dy : \mathbf{v_p} \longmapsto v_2,$$
$$dz : \mathbf{v_p} \longmapsto v_3.$$

Thus, dx, dy and dz behave with respect to tangent vectors in exactly the same way as x, y and z behave with respect to points.

You may wonder, with justification, what this use of dx has to do with its use in expressions such as

$$\int x^2\, dx.$$

The expression after the integral sign is a 1-form (it is df, where $f = x^3/3$). If finding directional derivatives of functions yields 1-forms, it is reasonable to think that integration might produce functions from 1-forms. Not only is this reasonable but, under suitable conditions, it is also true. Unfortunately, we do not have room to consider integration of 1-forms in this course, although *O'Neill* does mention the topic briefly in several places.

Example 5.3(2) A new idea is inserted here without explicit mention. So far, 1-forms act on tangent vectors (and produce real numbers), but we can form composites of 1-forms with vector fields. The case which is used in this example involves dx, etc. with the natural frame field U_1, etc.

When we write $dx(U_1)$, we mean the composite function

$$\mathbf{p} \longmapsto U_1(\mathbf{p}) \longmapsto dx(U_1(\mathbf{p})).$$

Since $U_1(\mathbf{p}) = (1,0,0)_{\mathbf{p}}$ and dx extracts the first component of a tangent vector, we have

$$dx(U_1(\mathbf{p})) = 1.$$

Similar calculations show that

$$\begin{array}{lll} dx(U_1(\mathbf{p})) = 1, & dx(U_2(\mathbf{p})) = 0, & dx(U_3(\mathbf{p})) = 0, \\ dy(U_1(\mathbf{p})) = 0, & dy(U_2(\mathbf{p})) = 1, & dy(U_3(\mathbf{p})) = 0, \\ dz(U_1(\mathbf{p})) = 0, & dz(U_2(\mathbf{p})) = 0, & dz(U_3(\mathbf{p})) = 1. \end{array}$$

These nine results can be summed up in 'functional' form as

$$dx_i(U_j) = \delta_{ij}.$$

The Kronecker delta, δ_{ij}, is a standard symbol and will be used a number of times in this course.

The definition of δ_{ij} is $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

If we use the linearity properties of 1-forms and the results summarized as

$$dx_i(U_j) = \delta_{ij},$$

then evaluating 1-forms on vector fields becomes quite routine.

Example Suppose that the vector field V and the 1-form ϕ are defined by

$$V = x^2U_1 - zU_2 + yU_3 \quad \text{and} \quad \phi = y\,dx + x\,dy - z\,dz.$$

Then

$$\begin{aligned} \phi(V) &= (y\,dx + x\,dy - z\,dz)(x^2U_1 - zU_2 + yU_3) \\ &= y\,dx(x^2U_1 - zU_2 + yU_3) + x\,dy(x^2U_1 - zU_2 + yU_3) \\ &\quad - z\,dz(x^2U_1 - zU_2 + yU_3) \quad \text{(linearity)} \\ &= y(x^2)\,dx(U_1) + y(-z)\,dx(U_2) + y(y)\,dx(U_3) + x(x^2)\,dy(U_1) \\ &\quad + x(-z)\,dy(U_2) + x(y)\,dy(U_3) - z(x^2)\,dz(U_1) - z(-z)\,dz(U_2) \\ &\quad - z(y)\,dz(U_3) \quad \text{(linearity)} \\ &= (x^2y + 0 + 0) + (0 - xz + 0) + (0 + 0 - zy) \quad (\text{using } dx_i(U_j) = \delta_{ij}) \\ &= x^2y - xz - yz. \end{aligned}$$

■

Having found that dx, dy and dz provide examples of 1-forms, *O'Neill* goes on to show that these three are all the 1-forms that you need (at least for $\mathbf{E}^3$). Lemma 5.4 shows that any 1-form can be expressed in terms of these 'basic' 1-forms.

Warning: Although differentials of functions provided the motivation for introducing 1-forms and will also provide many of the examples of 1-forms, it is quite easy to write down a 1-form which is not the differential of a function on $\mathbf{E}^3$. For example, the 1-form

$$dx + x\,dy + x^2\,dz$$

cannot be df for any f. To see why not, spend a few minutes trying to find f such that

$$\frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = x, \quad \frac{\partial f}{\partial z} = x^2.$$

Calculating differentials The remark after Lemma 5.7 is useful. Using

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz$$

is not usually the quickest way of finding df, as the example illustrates.

Also illustrated by the example is the fact that, now differentials are available, it is easier to find a directional derivative

$$\mathbf{v_p}[f]$$

by finding df and then using

$$\mathbf{v_p}[f] = df(\mathbf{v_p}).$$

Exercise 5.1 *O'Neill*, page 25, Exercise 1.

Exercise 5.2 *O'Neill*, page 25, Exercise 3.

Exercise 5.3 *O'Neill*, page 26, Exercise 4.

Exercise 5.4 *O'Neill*, page 26, Exercise 5 (first part only).

Exercise 5.5 *O'Neill*, page 26, Exercise 7.

[*Solutions on page 32*]

6 Differential forms

Read *O'Neill: Chapter I, Section 6, pages 26–31.*

Errata

1 *O'Neill*, page 31, Exercise 6, last line:

for $\ldots r, \phi, z \ldots$ read $\ldots r, \vartheta, z \ldots.$

2 *O'Neill*, page 29, line 8:

for $d\phi = d(xy)\,dx + d(x^2)\,dz.$ read $d\phi = d(xy) \wedge dx + d(x^2) \wedge dz.$

O'Neill is following his convention of omitting the wedge when one term is dx, dy or dz.

3 *O'Neill*, page 29, 4 lines from the bottom:

for $d\varphi$ read $d\phi$. ■

This section of *O'Neill* contains no difficult computational ideas. The problem is, rather, why introduce 2-forms and 3-forms at all? Little use is made of 3-forms in the course, but 2-forms appear in the work on surfaces.

2-forms Although it is not required for the exercises on this section, it seems reasonable to give some hints about what 2-forms *are*, in the same way as we have defined 1-forms to be real-valued functions on tangent vectors.

If you have retained a feeling, in spite of the last section, that dx has something to do with a 'small increment' in the x-direction, then you might suspect that the 2-form $dx\,dy$ might have something to do with a 'small area'. It is certainly the case that some of the 2-forms that we shall meet are connected with areas, although not all of them are.

Just as 1-forms operated on tangent vectors, 2-forms act on *pairs* of tangent vectors based at the same point. The details will emerge later in the course; for now we just give an example. The 2-form $dx\,dy$ acts on the pair of tangent vectors $(\mathbf{v_p}, \mathbf{w_p})$ at $\mathbf{p}$ as follows:

$$(dx\,dy)(\mathbf{v_p}, \mathbf{w_p}) = dx(\mathbf{v_p})\,dy(\mathbf{w_p}) - dx(\mathbf{w_p})\,dy(\mathbf{v_p}).$$

If we evaluate the terms, we obtain

$$(dx\,dy)(\mathbf{v_p}, \mathbf{w_p}) = v_1 w_2 - w_1 v_2.$$

The other 'basic' 2-forms $dx\,dz$ and $dy\,dz$ behave in a similar way.

The important thing at this stage is to be able to do calculations with differential forms. A more complete understanding of their meaning should emerge when we begin to use them later on.

Exterior derivative You might find it helpful to paraphrase Definition 6.3 as: 'to differentiate a 1-form, apply d to the coefficient functions'.

It is clearly possible to extend Definition 6.3 to 2-forms as follows.

Exterior derivative of 2-forms

If η is the 2-form

$$\eta = f_1\,dx\,dy + f_2\,dy\,dz + f_3\,dx\,dz,$$

then the exterior derivative of η is the 3-form

$$d\eta = df_1 \wedge dx\,dy + df_2 \wedge dy\,dz + df_3 \wedge dx\,dz.$$

Theorem 6.4 By now, the pattern used by *O'Neill* should be becoming familiar. Having defined a new form of differentiation (the exterior derivative), there follows a proof that the derivative obeys the linear and Leibniz rules that help in doing calculations. The point worthy of special note here is the form of Theorem 6.4(3) with its minus sign.

Optional Comment: If you have met the concepts of 'div', 'grad' and 'curl', you may be interested in page 31, Exercise 8, which explains the link between differential forms and these ideas.

For our purposes forms provide a very concise way of representing some of the ideas of differential geometry and we shall not use any of the other possible approaches from vector calculus.

Another, classical, approach uses tensors. We shall avoid them as well!

Exercise 6.1 *O'Neill*, page 31, Exercise 1.

Exercise 6.2 *O'Neill*, page 31, Exercise 2.

Exercise 6.3 *O'Neill*, page 31, Exercise 3.

Exercise 6.4 *O'Neill*, page 31, Exercise 4.

Exercise 6.5 *O'Neill*, page 31, Exercise 7. (*Hint*: Use the fact that any 1-form ϕ can be written as

$$\sum_{i=1}^{3} f_i \, dx_i$$

for suitable functions f_i.)

Exercise 6.6 *O'Neill*, page 32, Exercise 9.

[*Solutions on page 33*]

7 Mappings

Read *O'Neill: Chapter I, Section 7, pages 32–39.*

Errata

1 *O'Neill*, page 36, second line after the proof of Theorem 7.5:

for $F_*(v)$, read $F_*(\mathbf{v})$,

that is, the vector $\mathbf{v}$ should be in bold type.

2 *O'Neill*, page 37, proof of Corollary 7.7:

for 'in Corollary 7.6' read 'in Theorem 7.5'. ■

Mapping In this section, *O'Neill* generalizes the idea of a function

$$f : \mathbf{E}^3 \longrightarrow \mathbf{R}$$

to that of a function

$$F : \mathbf{E}^n \longrightarrow \mathbf{E}^m,$$

and if F is differentiable (in the sense defined on page 33) then such functions are called mappings. Thus, both real-valued functions on $\mathbf{E}^3$ and curves are special cases of mappings.

We shall use mappings for two purposes in this course. First, we shall use a special type of mapping (isometries) in our work on curves. Secondly, a mapping between a 'simple' surface, such as a sphere, and a more complicated one will help in the study of the shape of the more complicated surface. For this second purpose, we shall need to know how a mapping 'distorts' geometric objects (such as curves) in its domain.

Definition 7.1 This tells you how to specify a mapping: you give the coordinate functions.

If you inspect the paragraph following the definition carefully, you will see that the 'pointwise' and 'coordinate function' forms for specifying a mapping are very little different; the function form is more concise. You should find little difficulty in switching from one form to the other. What this paragraph does re-emphasize is that x, y and z are *functions*, whereas p_1, p_2 and p_3 are real numbers.

The final paragraph before Definition 7.2 makes an important point. It is not easy to decide at a glance what effect a mapping has on its domain. Just as functions from $\mathbf{R}$ to $\mathbf{R}$ can be investigated by using derivatives, so we shall use 'derivatives' of mappings to help in understanding them. To begin with, however, we look at what a mapping does to the route of a curve.

Example 7.3(1) In matrix form, with respect to the standard bases in domain and codomain, F can be written *M203*

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}.$$

Example 7.3(2) There is no reason why you should have thought of the idea of considering the image of the circle α. However, it does give some insight into what the mapping does.

Definition 7.4 This is the definition of a derivative for mappings mentioned above. If we write out explicitly the calculation involved, we can see that

$$F_*(\mathbf{v_p}) = (F(\mathbf{p} + t\mathbf{v}))'(0)_{F(\mathbf{p})}.$$

In this form, it is easier to see the family connection with the definition of directional derivative of a function on $\mathbf{E}^3$. Indeed, directional derivatives can now be seen as a special case of derivative maps.

It is worth emphasizing that

> F tells you what happens to each *point* $\mathbf{p}$ in the domain;
>
> F_* tells you what happens to each *tangent vector* $\mathbf{v_p}$ at $\mathbf{p}$,
>
> by mapping it to a tangent vector at $F(\mathbf{p})$.

A very simple example might help to explain why the derivative map can give information about the distorting properties of a mapping.

Example Consider the mapping

$$F : \mathbf{E}^2 \longrightarrow \mathbf{E}^2,$$
$$F = (x + y, y),$$

and the tangent vectors

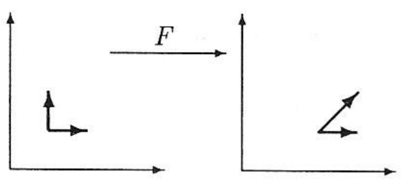
$$\mathbf{v_p} = (1, 0)_\mathbf{p},$$
$$\mathbf{w_p} = (0, 1)_\mathbf{p}.$$

Working from the definition,

$$\begin{aligned} F_*(\mathbf{v_p}) &= (F(\mathbf{p} + t(1,0)))'(0) \\ &= (F(p_1 + t, p_2))'(0) \\ &= (p_1 + t + p_2, p_2)'(0) \\ &= (1,0)(0) \\ &= (1,0) \quad \text{at } F(\mathbf{p}), \end{aligned}$$

$$\begin{aligned} F_*(\mathbf{w_p}) &= (F(\mathbf{p} + t(0,1)))'(0) \\ &= (F(p_1, p_2 + t))'(0) \\ &= (p_1 + p_2 + t, p_2 + t)'(0) \\ &= (1,1)(0) \\ &= (1,1) \quad \text{at } F(\mathbf{p}). \end{aligned}$$

Thus, at any point $\mathbf{p}$, the tangent vectors $(1,0)$ and $(0,1)$ are transformed to the tangent vectors $(1,0)$ and $(1,1)$ at the image $F(\mathbf{p})$.



This does give some idea of how the mapping distorts the domain. ■

Theorem 7.5 This result makes the connection between derivative maps and directional derivatives quite explicit. Paraphrasing: to find the value of the derivative map on $\mathbf{v_p}$, directionally differentiate the coordinate functions with respect to $\mathbf{v_p}$. (This paraphrase leaves out the point of application, but is a useful aid to memory nevertheless.)

The link with directional derivatives is the basis for a straightforward method of calculating derivative maps. The method for calculating directional derivatives in Lemma 3.2,

$$\mathbf{v_p}[f] = \frac{\partial f}{\partial x}(\mathbf{p})\, v_1 + \frac{\partial f}{\partial y}(\mathbf{p})\, v_2 + \frac{\partial f}{\partial z}(\mathbf{p})\, v_3,$$

can be written as a matrix product

$$\mathbf{v_p}[f] = \begin{pmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) & \frac{\partial f}{\partial y}(\mathbf{p}) & \frac{\partial f}{\partial z}(\mathbf{p}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

or, even,

$$\mathbf{v_p}[f] = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} (\mathbf{p}) \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Translating the result of Theorem 7.5 into matrix form, representing tangent vectors by column matrices with respect to the standard bases, leads directly to the idea of **Jacobian Matrix** discussed on page 37. For example, if

$$F = (f_1, f_2, f_3)$$

is a mapping from $\mathbf{E}^3$ to $\mathbf{E}^3$, then the image of a tangent vector $\mathbf{v_p}$ under F_* can be calculated as

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(\mathbf{p}) & \frac{\partial f_1}{\partial y}(\mathbf{p}) & \frac{\partial f_1}{\partial z}(\mathbf{p}) \\ \frac{\partial f_2}{\partial x}(\mathbf{p}) & \frac{\partial f_2}{\partial y}(\mathbf{p}) & \frac{\partial f_2}{\partial z}(\mathbf{p}) \\ \frac{\partial f_3}{\partial x}(\mathbf{p}) & \frac{\partial f_3}{\partial y}(\mathbf{p}) & \frac{\partial f_3}{\partial z}(\mathbf{p}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

at $F(\mathbf{p})$.

It is usual to shorten

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(\mathbf{p}) & \frac{\partial f_1}{\partial y}(\mathbf{p}) & \frac{\partial f_1}{\partial z}(\mathbf{p}) \\ \frac{\partial f_2}{\partial x}(\mathbf{p}) & \frac{\partial f_2}{\partial y}(\mathbf{p}) & \frac{\partial f_2}{\partial z}(\mathbf{p}) \\ \frac{\partial f_3}{\partial x}(\mathbf{p}) & \frac{\partial f_3}{\partial y}(\mathbf{p}) & \frac{\partial f_3}{\partial z}(\mathbf{p}) \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{pmatrix} (\mathbf{p}).$$

Strictly, the Jacobian matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{pmatrix}$$

represents F_* only with respect to the standard bases in domain and codomain.

However, it is an understandable and forgiveable abuse to write

$$F_* = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z} \end{pmatrix}$$

instead of 'F_* is represented by ... with respect to ...'.

Even more precisely, we should say that the matrix represents F_* only with respect to the natural frame field in domain and codomain. We should do so because F_* transforms tangent vectors, not vectors.

Linear approximation (Comment in second paragraph at top of page 37.) This comment refers to the link between the approximation

$$f(a+h) \approx f(a) + f'(a) \times h$$

for differentiable functions from $\mathbf{R}$ to $\mathbf{R}$ and the approximation

$$F(\mathbf{p}+\mathbf{v}) \approx F(\mathbf{p}) + F_*(\mathbf{v}).$$

The parallel is more obvious when the second approximation is written in column coordinate form:

$$F(\mathbf{p}+\mathbf{v}) \approx F(\mathbf{p}) + (\text{Jacobian matrix at } \mathbf{p}) \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Theorem 7.8 This result is important not only because it says that velocities of curves are preserved by F_* but also because it is the analogue of Lemma 4.6. It says that the straight line

$$\mathbf{p} + t\mathbf{v}$$

in the definition of F_* can be replaced by any curve through $\mathbf{p}$ with velocity $\mathbf{v}_\mathbf{p}$. Since F_* is obtained by directionally differentiating the coordinate functions, you should expect results that are true of directional derivatives to be true, in a suitable sense, of derivative maps.

Definition 7.9 and Theorem 7.10 These are mentioned here for completeness. In this course we shall use this definition and result in only one case. When we define surfaces, we shall use a two-dimensional version of the definition of curve. We shall have mappings from a suitable subset of $\mathbf{E}^2$ to $\mathbf{E}^3$. Although the surface defined by such mappings may be a very complicated object when viewed as a whole, we shall want any small piece of it to look quite like a piece of the plane. Regularity of the mappings used to define surfaces is exactly the condition that we need.

Checking regularity involves finding the rank of the Jacobian matrix. Now, the rank of a matrix is the maximum number of linearly independent rows (or columns). In this course, we are interested in mappings from $\mathbf{E}^n$ to $\mathbf{E}^m$ in the cases

It is a theorem of linear algebra that for *any* matrix the maximum number of linearly independent rows is the same as the maximum number of linearly independent columns.

$$n, m = 1, 2, 3.$$

If the dimension of the domain is greater than that of the codomain, then the mapping cannot be regular. For example, if

$$F : \mathbf{E}^3 \longrightarrow \mathbf{E}^2$$
$$F = (f_1, f_2),$$

then the Jacobian matrix representing F_* is

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \end{pmatrix} (\mathbf{p}).$$

This cannot have more than 2 independent rows, so its rank cannot exceed 2, which is less than the dimension of the domain.

In practice, the commonest situation in which you will be required to check for regularity is for mappings

$$\mathbf{E}^2 \longrightarrow \mathbf{E}^3,$$

where the task reduces to showing that the 3×2 Jacobian matrix has two independent columns.

Composite mappings An important result is hidden by *O'Neill* in two exercises. Because of the importance of the result, we ask you to do the two exercises now. Our solutions (with comments) follow immediately.

Try *O'Neill*, page 40, Exercises 7 and 12(b) now.

O'Neill writes GF for the composite, where you may be used to writing $G \circ F$.

Solution to Exercise 7 In full:

$$\begin{aligned} G(F(p_1, p_2)) &= G(f_1(p_1, p_2), f_2(p_1, p_2)) \\ &= (g_1(f_1(p_1, p_2), f_2(p_1, p_2)), g_2(f_1(p_1, p_2), f_2(p_1, p_2))). \end{aligned}$$

In function form:

$$GF = (g_1(f_1, f_2), g_2(f_1, f_2)).$$

Because F and G are mappings, their coordinate functions are differentiable. Thus, the coordinate functions of GF involve composites of differentiable functions. Hence, the coordinate functions of GF are differentiable and GF is a mapping.

Solution to Exercise 12(b) This requires three applications of Theorem 7.8. First, let

$$\beta = G(F(\alpha)).$$

Then, by Theorem 7.8 applied to GF

$$(GF)_*(\alpha') = \beta'.$$

Now, if we define

$$\gamma = F(\alpha)$$

and apply Theorem 7.8 to F, we obtain

$$F_*(\alpha') = \gamma'.$$

But

$$\beta = G(\gamma),$$

so

$$G_*(\gamma') = \beta'.$$

Putting these results together gives

$$\begin{aligned} (GF)_*(\alpha') &= \beta' \\ &= G_*(\gamma') \\ &= G_*(F_*(\alpha')). \end{aligned}$$

The consequence of this last result is a 'composite rule' for derivative maps. Since

$$(GF)_*(\alpha') = G_*(F_*(\alpha'))$$

holds for *all* possible curves in the domain, we conclude that

$$(GF)_* = G_*(F_*).$$

The composite rule as stated above is a little dangerous when using Jacobian matrices to carry out calculations. If we recall the composite rule for functions from $\mathbf{R}$ to $\mathbf{R}$ in the form

$$(g \circ f)' = g'(f) \times f',$$

perhaps you can see why. The derivative g' must be evaluated at the image under the first function, f.

Consider what happens under the various composites that we are discussing. The mappings transform points:

$$\mathbf{p} \longmapsto F(\mathbf{p}) \longmapsto G(F(\mathbf{p})).$$

The derivative maps transform tangent vectors:

$$\mathbf{v_p} \longmapsto F_*(\mathbf{v_p})_{F(\mathbf{p})} \longmapsto (GF)_*(\mathbf{v_p})_{G(F(\mathbf{p}))}.$$

It should be clear that F_* can be achieved by applying the value of the Jacobian matrix for F at $\mathbf{p}$ but that G_* requires the value of the Jacobian matrix for G at $F(\mathbf{p})$.

This last statement suggests that a 'safer' statement of the composite rule for mappings is

$$(G \circ F)_* = G_*(F)F_*,$$

where $G_*(F)$ means G_* evaluated at F.

This alternative form of the composite rule for maps has the advantage of being very similar to the one for ordinary functions.

The similarity may make it more memorable.

The set of exercises for this section includes practice at finding derived maps of composites using Jacobians.

Exercise 7.1 Express each of the following mappings from $\mathbf{E}^3$ to $\mathbf{E}^3$ in terms of the Euclidean coordinate functions.

(a) $F : \mathbf{p} \longmapsto -3\mathbf{p}$ (b) $F : \mathbf{p} \longmapsto (e^{p_1 p_2}, p_3 + 2p_2, p_1^2)$

Exercise 7.2 The mapping $F : \mathbf{E}^2 \longrightarrow \mathbf{E}^2$ is defined by

$$F = (2xy, x^2 - y^2).$$

Find the images under F of the following lines, and identify the image curves.

(a) $x = 1$ (b) $y = -1$

Exercise 7.3 Let $F : \mathbf{E}^3 \longrightarrow \mathbf{E}^3$ be the mapping

$$F = (x - y, x + y, 2z),$$

from *O'Neill*, Example 7.3.

(a) Find, from first principles, the derivative map F_*.

(b) Show that F is linear, in the sense that

$$F(a\mathbf{p} + b\mathbf{q}) = aF(\mathbf{p}) + bF(\mathbf{q}), \quad \text{for all } \mathbf{p}, \mathbf{q} \in \mathbf{E}^3,\ a, b \in \mathbf{R}.$$

(c) Generalize the result for F to show that if G is *any* linear transformation (mapping) from $\mathbf{E}^3$ to $\mathbf{E}^3$, then

$$G_*(\mathbf{v_p}) = G(\mathbf{v})_{G(\mathbf{p})}.$$

Exercise 7.4 Let F be a mapping from $\mathbf{E}^3$ to $\mathbf{E}^3$ and let G be a mapping from $\mathbf{E}^3$ to $\mathbf{E}^2$ defined by

$$F = (x + y, x - 2y, x + y + z) \quad \text{and} \quad G = (x^2 + y^2, y^2 + z^2).$$

(a) Write down the Jacobian matrix for F_*.

(b) Write down the Jacobian matrix for G_* and evaluate it at $F(x, y, z)$.

(c) Use the composite rule to find the Jacobian matrix for $(G \circ F)_*$.

[*Solutions on page 35*]

8 Summary

Read *O'Neill: Chapter I, Section 8, page 41.*

The first chapter of *O'Neill* has not provided all the tools needed for our study of curves. We shall need a little more discussion of frame fields and some work on reparametrization. The first two sections of *Part II* will provide what is necessary.

There is one 'spin-off' from the work on derivative maps that we have done. We can provide a general version of the chain rule. Suppose that we extend the work done on Exercise 7, page 40 of *O'Neill*.

We have mappings

$$F = (f_1, f_2), \quad G = (g_1, g_2)$$

from $\mathbf{E}^2$ to $\mathbf{E}^2$. The corresponding Jacobians are

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} \end{pmatrix}.$$

The composite mapping $G \circ F$ has Jacobian

$$\begin{pmatrix} \dfrac{\partial g_1(f_1, f_2)}{\partial x} & \dfrac{\partial g_1(f_1, f_2)}{\partial y} \\ \dfrac{\partial g_2(f_1, f_2)}{\partial x} & \dfrac{\partial g_2(f_1, f_2)}{\partial y} \end{pmatrix}.$$

Writing out the composite rule for derived maps gives

$$\begin{pmatrix} \dfrac{\partial g_1(f_1, f_2)}{\partial x} & \dfrac{\partial g_1(f_1, f_2)}{\partial y} \\ \dfrac{\partial g_2(f_1, f_2)}{\partial x} & \dfrac{\partial g_2(f_1, f_2)}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} \end{pmatrix} (F) \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix}.$$

Taking the top left element of both sides gives

$$\frac{\partial g_1(f_1, f_2)}{\partial x} = \frac{\partial g_1}{\partial x}(f_1, f_2)\frac{\partial f_1}{\partial x} + \frac{\partial g_1}{\partial y}(f_1, f_2)\frac{\partial f_2}{\partial x}.$$

There are three other, similar results.

A similar approach for a function of three variables would give

$$\frac{\partial g_1(f_1, f_2, f_3)}{\partial x} = \frac{\partial g_1}{\partial x}(f_1, f_2, f_3)\frac{\partial f_1}{\partial x} + \frac{\partial g_1}{\partial y}(f_1, f_2, f_3)\frac{\partial f_2}{\partial x} + \frac{\partial g_1}{\partial z}(f_1, f_2, f_3)\frac{\partial f_3}{\partial x}.$$

You might like to show that the composite rule for derived maps reduces to the ordinary chain rule in the case of mappings (functions) from $\mathbf{R}$ to $\mathbf{R}$. (*Hint*: A 1×1 Jacobian matrix is just like a function.)

Solutions to the exercises

Solution 1.1

(a) We have

$$fg^2 = (x^2y)(y\sin z)^2 = x^2y^3\sin^2 z.$$

(b) Here

$$\frac{\partial f}{\partial x} = 2xy,$$

$$\frac{\partial g}{\partial y} = \sin z,$$

$$\begin{aligned}\frac{\partial f}{\partial x}g + \frac{\partial g}{\partial y}f &= (2xy)(y\sin z) + (\sin z)(x^2y)\\ &= xy(2y+x)\sin z.\end{aligned}$$

(c) Note that, by definition,

$$\frac{\partial^2(fg)}{\partial y\partial z} = \frac{\partial}{\partial y}\frac{\partial(fg)}{\partial z}.$$

We differentiate first with respect to z, then with respect to y. Although it is perfectly easy to find fg and work on that, we shall indicate how to use the product rule:

$$\begin{aligned}\frac{\partial(fg)}{\partial z} &= \frac{\partial f}{\partial z}g + f\frac{\partial g}{\partial z}\\ &= 0\times g + f\times y\cos z\\ &= x^2y^2\cos z,\\ \frac{\partial^2(fg)}{\partial y\partial z} &= \frac{\partial(x^2y^2\cos z)}{\partial y}\\ &= 2x^2y\cos z.\end{aligned}$$

(d) This provides an example suitable for using the chain (composite) rule:

$$\begin{aligned}\frac{\partial\sin(f)}{\partial y} &= \frac{d(\sin(f))}{df}\frac{\partial f}{\partial y}\\ &= \cos(f)\,x^2\\ &= x^2\cos(x^2y).\end{aligned}$$

Note that we could have written the start of this solution as

$$\ldots\sin'(f)\frac{\partial f}{\partial y} = \ldots.$$

Solution 1.2

(a) This solution is full to the point of tediousness just to illustrate what is meant by the pointwise principle (abbreviated here to pp).

$$\begin{aligned}f(1,1,1) &= (x^2y - y^2z)(1,1,1)\\ &= (x^2y)(1,1,1) - (y^2z)(1,1,1) \quad \text{(pp for subtraction)}\\ &= (x^2)(1,1,1)y(1,1,1) - (y^2)(1,1,1)z(1,1,1)\\ &\qquad \text{(pp for multiplication)}\\ &= x(1,1,1)x(1,1,1)y(1,1,1) - y(1,1,1)y(1,1,1)z(1,1,1)\\ &\qquad \text{(pp for multiplication)}\\ &= 1\times1\times1 - 1\times1\times1 = 0.\end{aligned}$$

In practice, you probably did the calculation with a fraction of the effort expended here.

(b) This time we give a shorter solution!

$$\begin{aligned}f(3,-1,1/2) &= (x^2y - y^2z)(3,-1,1/2)\\ &= 3^2\times(-1) - (-1)^2\times(1/2)\\ &= -9 - 1/2 = -19/2.\end{aligned}$$

(c) Here

$$f(a,1,1-a) = a^2\times1 - 1^2\times(1-a) = a^2 + a - 1.$$

(d) This time

$$f(t,t^2,t^3) = t^2\times t^2 - (t^2)^2t^3 = t^4 - t^7.$$

Solution 1.3

These solutions use the linearity properties, the Leibniz property and the chain rule as appropriate.

(a) Here

$$\begin{aligned}\frac{\partial f}{\partial x} &= \sin(xy) + x\frac{\partial\sin(xy)}{\partial x} + 0\cos(xz) + y\frac{\partial\cos(xz)}{\partial x}\\ &= \sin(xy) + xy\cos(xy) - yz\sin(xz).\end{aligned}$$

(b) In this case

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{d(\sin(g))}{dg}\frac{\partial g}{\partial x} \quad \text{(chain rule)}\\ &= \cos g\frac{d(e^h)}{dh}\frac{\partial h}{\partial x} \quad \text{(chain rule)}\\ &= \cos g\, e^h\, 2x\\ &= 2xe^{x^2+y^2+z^2}\cos\left(e^{x^2+y^2+z^2}\right).\end{aligned}$$

Solution 1.4

What would be extremely useful for this question is a generalized chain rule. However, for now, we calculate the composite function in each case and differentiate directly.

(a) Here

$$f = (x+y)^2 - (y^2)(x+z),$$

so

$$\frac{\partial f}{\partial x} = 2(x+y) - y^2(1+0) = 2(x+y) - y^2.$$

(b) This time

$$\begin{aligned}f &= e^{2x} - e^{x+y}e^x\\ &= e^{2x} - e^{2x}e^y\\ &= e^{2x}(1-e^y),\end{aligned}$$

so

$$\frac{\partial f}{\partial x} = 2e^{2x}(1-e^y).$$

(c) Now

$$f = x^2 - (-x)(x) = 2x^2,$$

so

$$\frac{\partial f}{\partial x} = 4x.$$

Solution 2.1

By definition, the vector part of $3\mathbf{v_p} - 2\mathbf{w_p}$ is

$$\begin{aligned}3\mathbf{v} - 2\mathbf{w} &= 3(-2,1,-1) - 2(0,1,3)\\ &= (-6,1,-9).\end{aligned}$$

Hence

$$3\mathbf{v_p} - 2\mathbf{w_p} = (-6,1,-9)_\mathbf{p}.$$

This result could also be written

$$3\mathbf{v_p} - 2\mathbf{w_p} = -6U_1(\mathbf{p}) + U_2(\mathbf{p}) - 9U_3(\mathbf{p}).$$

Solution 2.2

First, we find $W - xV$;

$$\begin{aligned} W - xV &= (2x^2U_2 - U_3) - x(xU_1 + yU_2) \\ &= 2x^2U_2 - U_3 - x^2U_1 - xyU_2 \\ &= -x^2U_1 + (2x^2 - xy)U_2 - U_3. \end{aligned}$$

Now, the value at $\mathbf{p} = (-1, 0, 2)$ of this vector field is

$$\begin{aligned} &(-x^2U_1 + (2x^2 - xy)U_2 - U_3)(-1,0,2) \\ = &-(-1)^2U_1(-1,0,2) + (2(-1)^2 - (-1)(0))U_2(-1,0,2) \\ &- U_3(-1,0,2) \\ = &-U_1(-1,0,2) + 2U_2(-1,0,2) - U_3(-1,0,2) \\ = &-(1,0,0)_{(-1,0,2)} + 2(0,1,0)_{(-1,0,2)} - (0,0,1)_{(-1,0,2)} \\ = &(-1,2,-1)_{(-1,0,2)}. \end{aligned}$$

Solution 2.3

(a) Since

$$7V = 2z^2U_1 - xyU_3,$$

we have

$$V = \frac{2z^2}{7}U_1 + 0U_2 - \frac{xy}{7}U_3.$$

You may, quite reasonably, have omitted the middle term since it is zero.

(b) We can rewrite the information given as

$$V(\mathbf{p}) = p_1U_1(\mathbf{p}) + (p_3 - p_1)U_2(\mathbf{p}) + 0U_3(\mathbf{p}).$$

If we also use $x(\mathbf{p}) = p_1$, etc., we obtain

$$V(\mathbf{p}) = x(\mathbf{p})U_1(\mathbf{p}) + (z(\mathbf{p}) - x(\mathbf{p}))U_2(\mathbf{p}) + 0U_3(\mathbf{p})$$

which, by the pointwise principle, defines V to be

$$V = xU_1 + (z - x)U_2 + 0U_3.$$

(c) We have $V = xU_1 + 2yU_2 + xy^2U_3$.

(d) First, we need the vector part of $V(\mathbf{p})$, which is

$$(1 + p_1, p_2p_3, p_2) - (p_1, p_2, p_3) = (1, p_2p_3 - p_2, p_2 - p_3).$$

By an analysis like that in the first part of this question, we obtain

$$V = U_1 + y(z-1)U_2 + (y - z)U_3.$$

(e) Here the vector part is $(-p_1, -p_2, -p_3)$, so

$$V(\mathbf{p}) = (-p_1, -p_2, -p_3)_{\mathbf{p}}$$

and so

$$V = -xU_1 - yU_2 - zU_3.$$

Solution 2.4

(a) By the definition of independence of tangent vectors, it is enough to show that the vector parts are linearly independent. The vector parts are

$$(1, 0, -p_1), \quad (0, 1, 0) \quad \text{and} \quad (p_1, 0, 1).$$

We apply the standard test for linear independence. Suppose that

$$a(1,0,-p_1) + b(0,1,0) + c(p_1,0,1) = \mathbf{0}, \quad a, b, c \in \mathbf{R}.$$

Comparing components gives

$$\begin{aligned} a + cp_1 &= 0, \\ b &= 0, \\ -ap_1 + c &= 0. \end{aligned}$$

We have $b = 0$ immediately. The third equation gives $c = ap_1$, and substituting in the first gives

$$a + ap_1^2 = a(1 + p_1^2) = 0.$$

Since $1 + p_1^2 > 0$, we have $a = 0$ and hence $c = 0$. Thus

$$a = b = c = 0,$$

and the three vectors are linearly independent.

(b) We want

$$\begin{aligned} xU_1 + yU_2 + zU_3 &= fV_1 + gV_2 + hV_3 \\ &= f(U_1 - xU_3) + gU_2 + h(xU_1 + U_3), \end{aligned}$$

where f, g and h are functions from $\mathbf{E}^3$ to $\mathbf{R}$. Comparing the coefficients of the U_i gives

$$\begin{aligned} x &= f + hx, \\ y &= g, \\ z &= -fx + h. \end{aligned}$$

We have

$$g = y,$$

and substituting $f = x - hx$ from the first equation into the third gives

$$\begin{aligned} z &= -(x - hx)x + h \\ &= h(1 + x^2) - x^2. \end{aligned}$$

Hence

$$h = (x^2 + z)/(1 + x^2)$$

and

$$\begin{aligned} f &= x - \frac{x^2 + z}{1 + x^2}x \\ &= x\left(1 - \frac{x^2 + z}{1 + x^2}\right) \\ &= \frac{x(1 - z)}{1 + x^2}. \end{aligned}$$

Hence

$$xU_1 + yU_2 + zU_3 = \frac{x(1-z)}{1+x^2}V_1 + yV_2 + \frac{x + z}{1 + x^2}V_3.$$

Solution 3.1

To apply the definition we shall need $\mathbf{p} + t\mathbf{v}$:

$$\mathbf{p} + t\mathbf{v} = (2 + 2t, -t, -1 + 3t).$$

(a) Using the above:

$$\begin{aligned} \mathbf{v_p}[f] &= ((y^2z)(2 + 2t, -t, -1 + 3t))'(0) \\ &= ((-t)^2(-1 + 3t))'(0) \\ &= (2t(-1 + 3t) + t^2(3))(0) \quad \text{(product rule)} \\ &= 0 + 0 = 0. \end{aligned}$$

(b) Working as in the last part,

$$\begin{aligned} \mathbf{v_p}[f] &= ((2 + 2t)^7)'(0) \\ &= (7(2 + 2t)^6 \times 2)(0) \\ &= 7 \times 2^6 \times 2 \\ &= 896. \end{aligned}$$

(c) This time

$$\begin{aligned} \mathbf{v_p}[f] &= (e^{2+2t}\cos(-t))'(0) \\ &= (2e^{2+2t}\cos(-t) + e^{2+2t}(-\sin(-t))(-1))(0) \\ &= 2e^2\cos 0 + e^2 \sin 0 \\ &= 2e^2. \end{aligned}$$

Solution 3.2

For each part, we calculate the three partial derivatives, evaluate at $\mathbf{p}$ and then apply the formula from Lemma 3.2.

(a) $$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial x}(\mathbf{p}) = 0,$$
$$\frac{\partial f}{\partial y} = 2yz, \quad \frac{\partial f}{\partial y}(\mathbf{p}) = 2(0)(-1) = 0,$$
$$\frac{\partial f}{\partial z} = y^2, \quad \frac{\partial f}{\partial z}(\mathbf{p}) = 0^2 = 0,$$
$$\mathbf{v}_\mathbf{p}[f] = 0 \times 2 + 0 \times (-1) + 0 \times 3 = 0.$$

(b) $$\frac{\partial f}{\partial x} = 7x^6, \quad \frac{\partial f}{\partial x}(\mathbf{p}) = 7 \times 2^6 = 448,$$
$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial y}(\mathbf{p}) = 0,$$
$$\frac{\partial f}{\partial z} = 0, \quad \frac{\partial f}{\partial z}(\mathbf{p}) = 0,$$
$$\mathbf{v}_\mathbf{p}[f] = 448(2) + 0 + 0 = 896.$$

(c) $$\frac{\partial f}{\partial x} = e^x \cos y, \quad \frac{\partial f}{\partial x}(\mathbf{p}) = e^2 \cos 0 = e^2,$$
$$\frac{\partial f}{\partial y} = -e^x \sin y, \quad \frac{\partial f}{\partial y}(\mathbf{p}) = -e^2 \sin 0 = 0,$$
$$\frac{\partial f}{\partial z} = 0, \quad \frac{\partial f}{\partial z}(\mathbf{p}) = 0,$$
$$\mathbf{v}_\mathbf{p}[f] = e^2(2) + 0 + 0 = 2e^2.$$

Solution 3.3

The main technique used here is $U_i[f] = \partial f/\partial x_i$. We also use linearity but, as usual, this is more or less taken for granted.

$$\begin{aligned}
V[f] &= (2U_1 - U_2 + 3U_3)[f] \\
&= 2U_1[f] - U_2[f] + 3U_3[f] \quad \text{(linearity)} \\
&= 2\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} + 3\frac{\partial f}{\partial z} \\
&= 2e^x \cos y - e^x(-\sin y) + 3(0) \\
&= e^x(2\cos y + \sin y).
\end{aligned}$$

Solution 3.4

These questions are all tackled by applying the various parts of Corollary 3.4 (pages 13–14) and $U_i[f] = \partial f/\partial x_i$.

(a) Here
$$\begin{aligned}
V[f] &= (y^2U_1 - xU_3)[f] \\
&= y^2U_1[f] - xU_3[f] \quad \text{(Corollary 3.4(1))} \\
&= y^2\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial z} \\
&= y^2(y) - x(0) \\
&= y^3.
\end{aligned}$$

(b) This time,
$$\begin{aligned}
V[g] &= y^2U_1[g] - xU_3[g] \\
&= y^2\frac{\partial g}{\partial x} - x\frac{\partial g}{\partial z} \\
&= y^2(0) - x(3z^2) \\
&= -3xz^2.
\end{aligned}$$

(c) Now
$$\begin{aligned}
V[fg] &= V[f]\,g + f\,V[g] \quad \text{(Corollary 3.4(3))} \\
&= y^3g + f\,(-3xz^2) \quad \text{(using parts (i) and (ii))} \\
&= y^3(z^3) + xy(-3xz^2) \\
&= yz^2(y^2z - 3x^2).
\end{aligned}$$

(d) In this case,
$$\begin{aligned}
fV[g] - gV[f] &= xy(-3xz^2) - z^3(y^3) \\
&= -yz^2(3x^2 + y^2z).
\end{aligned}$$

(e) We could do this question very simply by calculating $f^2 + g^2$.

However, there is some merit in seeing how Corollary 3.4 gives an extension to a familiar result from ordinary calculus. First, note
$$\begin{aligned}
V[f^2] &= V[ff] \\
&= V[f]\,f + f\,V[f] \\
&= 2f\,V[f].
\end{aligned}$$
This result may remind you of applying the (ordinary) chain rule to obtain
$$(f^2)' = 2ff'.$$
Applying the result above and a corresponding one for $V[g^2]$,
$$\begin{aligned}
V[f^2 + g^2] &= V[f^2] + V[g^2] \quad \text{(linearity)} \\
&= 2f\,V[f] + 2g\,V[g] \\
&= 2(xy)(y^3) + 2(z^3)(-3xz^2) \\
&= 2x(y^4 - 3z^5).
\end{aligned}$$

(f) Now, we have
$$\begin{aligned}
V[V[f]] &= V[y^3] \\
&= (y^2U_1 - xU_3)[y^3] \\
&= y^2U_1[y^3] - xU_3[y^3] \\
&= y^2\frac{\partial y^3}{\partial x} - x\frac{\partial y^3}{\partial z} \\
&= y^2(0) - x(0) \\
&= 0.
\end{aligned}$$

Solution 3.5

The question directs us to calculate the three derivatives
$$V[x_i] = v_1U_1[x_i] + v_2U_2[x_i] + v_3U_3[x_i], \quad i = 1, 2, 3.$$
To do this, note that
$$\frac{\partial x_1}{\partial x_1} = 1, \quad \frac{\partial x_1}{\partial x_2} = 0, \quad \text{etc.}$$
That is,
$$\frac{\partial x_i}{\partial x_i} = 1, \quad \frac{\partial x_i}{\partial x_j} = 0, \quad i \neq j.$$
Hence
$$V[x_i] = v_1U_1[x_i] + v_2U_2[x_i] + v_3U_3[x_i] = v_i, \quad i = 1, 2, 3.$$
This is because only the term whose suffix is i is non-zero. Hence, by the definition of V,
$$\begin{aligned}
V &= v_1U_1 + v_2U_2 + v_3U_3 \\
&= V[x_1]U_1 + V[x_2]U_2 + V[x_3]U_3 \\
&= \sum_{i=1}^{3} V[x_i]U_i.
\end{aligned}$$

Solution 3.6

The method of solving this question illustrates a good general point. When some property is given to hold for *all* of some class of objects, it is always worth applying the information to some 'easy' particular cases. Here we are told that $V[f] = W[f]$ for all real-valued functions f on $\mathbf{E}^3$. We may obtain useful information by applying

this to the coordinate functions x_i, $i = 1, 2, 3$. We shall use the standard convention that

$$V = v_1U_1 + v_2U_2 + v_3U_3$$
$$W = w_1U_1 + w_2U_2 + w_3U_3.$$

By what we know,

$$V[x_i] = W[x_i], \quad i = 1, 2, 3,$$

but we also know that

$$V[x_i] = v_i, \; W[x_i] = w_i,$$

so

$$v_i = w_i, \quad i = 1, 2, 3.$$

This means that $V = W$.

Solution 4.1

(a) We obtain the velocity by differentiating each coordinate function:

$$\alpha'(t) = (-4\cos t \sin t, 2\cos 2t, 2\cos t)_{\alpha(t)}, \quad 0 < t < \frac{\pi}{2}.$$

Note that the velocity vector has a point of application since it is a tangent vector (as well as being a vector tangent to the curve).

Evaluating at $\pi/4$, we use

$$\sin\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}},$$
$$\cos\frac{\pi}{2} = 0, \quad \sin\frac{\pi}{2} = 1.$$

Thus

$$\alpha'\left(\frac{\pi}{4}\right) = (-2, 0, \sqrt{2})_{(1,1,\sqrt{2})}.$$

(b) (i) The length of $\alpha(t)$ is given by

$$\begin{aligned}\|\alpha(t)\| &= \sqrt{4\cos^4 t + \sin^2 2t + 4\sin^2 t}\\ &= \sqrt{4\cos^4 t + 4\sin^2 t\cos^2 t + 4\sin^2 t}\\ &\quad (\sin 2t = 2\sin t\cos t)\\ &= \sqrt{4\cos^2 t(\cos^2 t + \sin^2 t) + 4\sin^2 t}\\ &= \sqrt{4\cos^2 t + 4\sin^2 t}\\ &= \sqrt{4} = 2.\end{aligned}$$

(ii) By similar methods,

$$\begin{aligned}\|\alpha(t) - (1, 0, 2\sin t)\| &= \sqrt{(2\cos^2 t - 1)^2 + \sin^2 2t + 0^2}\\ &= \sqrt{\cos^2 2t + \sin^2 2t}\\ &\quad (\cos 2t = 2\cos^2 t - 1)\\ &= \sqrt{1} = 1.\end{aligned}$$

(c) Given that

$$t = h(s) = \arcsin s,$$

that is,

$$s = \sin t,$$

we shall need $\cos t$ in terms of s as well:

$$\cos t = \sqrt{1 - \sin^2 t} = \sqrt{1 - s^2}.$$

Hence, using

$$\sin 2t = 2\sin t\cos t,$$

we have

$$\begin{aligned}\beta(s) &= \alpha(h(s))\\ &= (2(\sqrt{1-s^2})^2, 2s\sqrt{1-s^2}, 2s)\\ &= (2(1-s^2), 2s\sqrt{1-s^2}, 2s).\end{aligned}$$

Solution 4.2

If we assume that the coordinate functions of α are α_1, α_2 and α_3, then we know that

$$\alpha_1(0) = 1, \quad \alpha_1'(t) = t^2,$$
$$\alpha_2(0) = 0, \quad \alpha_2'(t) = t,$$
$$\alpha_3(0) = -5, \quad \alpha_3'(t) = e^t.$$

Integrating, and using the initial values, gives

$$\alpha_1(t) = \frac{t^3}{3} + 1,$$
$$\alpha_2(t) = \frac{t^2}{2},$$
$$\alpha_3(t) = e^t - 5,$$

so

$$\alpha(t) = \left(\frac{t^3}{3} + 1, \frac{t^2}{2}, e^t - 5\right).$$

Solution 4.3

We begin by expressing the equation of each line in

$$\mathbf{p} + t\mathbf{v}$$

form. For the first line,

$$\mathbf{p} = (1, -3, -1),$$
$$\mathbf{v} = (6, 2, 1) - (1, -3, -1) = (5, 5, 2).$$

Hence the equation is

$$\begin{aligned}\alpha(t) &= (1, -3, -1) + t(5, 5, 2)\\ &= (1 + 5t, -3 + 5t, -1 + 2t).\end{aligned}$$

For the second line,

$$\mathbf{p} = (-1, 1, 0),$$
$$\mathbf{v} = (-5, -1, -1) - (-1, 1, 0) = (-4, -2, -1).$$

This gives

$$\beta(s) = (-1 - 4s, 1 - 2s, -s).$$

In order to discover if these lines meet, we must try to solve

$$(1 + 5t, -3 + 5t, -1 + 2t) = (-1 - 4s, 1 - 2s, -s).$$

Comparing the third components shows that, if there is a solution, we must have $s = 1 - 2t$. The first component now gives

$$1 + 5t = -1 - 4s = -1 - 4(1 - 2t) = 8t - 5,$$

so

$$6 = 3t$$

or

$$t = 2.$$

Hence

$$s = 1 - 4 = -3.$$

Checking the second components:

$$-3 + 5t = -3 + 5(2) = 7,$$
$$1 - 2s = 1 - (-6) = 7.$$

Thus, taking $t = 2$, $s = -3$ shows that the point $(11, 7, 3)$ lies on both lines.

Solution 4.4

(a) The three velocity vectors are

$$(1, 2t, 1)_{(t, 1+t^2, t)},$$
$$(\cos t, -\sin t, 1)_{(\sin t, \cos t, t)} \quad \text{and}$$
$$(\cosh t, \sinh t, 1)_{(\sinh t, \cosh t, t)}.$$

It follows that the velocity vectors at $t = 0$ are all

$$\mathbf{v}_{\mathbf{P}} = (1, 0, 1)_{(0,1,0)}.$$

(b) We apply the corollary to Lemma 4.6 that was discussed earlier in this section. For each curve, we evaluate f on the curve, differentiate and then substitute $t = 0$:

$$f(t, 1+t^2, t) = t^2 - (1+t^2)^2 + t^2,$$
$$\text{Derivative} = 2t - 2(1+t^2)(2t) + 2t,$$
$$\text{At } t = 0, \text{ value} = 0 - 0 + 0.$$

Hence $\mathbf{v}_{\mathbf{P}}[f] = 0$.

Similarly with the second curve:

$$f(\sin t, \cos t, t) = \sin^2 t - \cos^2 t + t^2,$$
$$\text{Derivative} = 2\sin t \cos t + 2\cos t \sin t + 2t,$$
$$\text{At } t = 0 \text{ value} = 0 + 0 + 0 = 0.$$

For the third:

$$f(\sinh t, \cosh t, t) = \sinh^2 t - \cosh^2 t + t^2,$$
$$\text{Derivative} = 2\sinh t \cosh t - 2\cosh t \sinh t + 2t,$$
$$\text{At } t = 0 \text{ value} = 0 + 0 + 0 = 0.$$

Solution 4.5

Following the definition given in the question, we first find the velocity vector:

$$\alpha'(t) = (-2\sin t, 2\cos t, 1)_{(2\cos t, 2\sin t, t)}.$$

Now we write down the tangent line:

$$u \longmapsto (2\cos t, 2\sin t, t) + u(-2\sin t, 2\cos t, 1),$$

that is,

$$u \longmapsto (2(\cos t - u\sin t), 2(\sin t + u\cos t), t + u).$$

Hence, at $t = 0$ we obtain

$$u \longmapsto (2, 2u, u).$$

At $t = \pi/4$, the tangent line is

$$u \longmapsto (\sqrt{2}(1-u), \sqrt{2}(1+u), \pi/4 + u).$$

Note: In the above solution we have opted to find the general form of the tangent line and then substitute to obtain the two special cases asked for. You may well have dealt with each special case individually.

Solution 5.1

(a) We use the properties of dx, etc., that is

$$dx(\mathbf{v}_{\mathbf{P}}) = v_1, \quad dy(\mathbf{v}_{\mathbf{P}}) = v_2, \quad dz(\mathbf{v}_{\mathbf{P}}) = v_3.$$

We have

$$\begin{aligned}(y^2\,dx)((1,2,-3)_{(0,-2,1)}) &= y^2(0,-2,1)\,dx(1,2,-3)\\ &= (-2)^2(1)\\ &= 4.\end{aligned}$$

(b) Similarly,

$$\begin{aligned}&(z\,dy - y\,dz)((1,2,-3)_{(0,-2,1)})\\ &= z(0,-2,1)dy(1,2,-3) - y(0,-2,1)dz(1,2,-3)\\ &= 1(2) - (-2)(-3)\\ &= -4.\end{aligned}$$

(c) Here,

$$\begin{aligned}&((z^2-1)dx - dy + x^2\,dz)((1,2,-3)_{(0,-2,1)})\\ &= (1^2-1)(1) - (2) + 0^2(-3)\\ &= 0 - 2 + 0\\ &= -2.\end{aligned}$$

Solution 5.2

In this solution we make heavy use of the values of dx, dy and dz on U_1, U_2 and U_3:

$$\begin{aligned}dx(U_1) &= 1, & dx(U_2) &= 0, & dx(U_3) &= 0,\\ dy(U_1) &= 0, & dy(U_2) &= 1, & dy(U_3) &= 0,\\ dz(U_1) &= 0, & dz(U_2) &= 0, & dz(U_3) &= 1.\end{aligned}$$

(a) Expanding ϕ and using linearity gives

$$\begin{aligned}\phi(V) &= (x^2\,dx - y^2\,dz)(xU_1 + yU_2 + zU_3)\\ &= x^2\,dx(xU_1 + yU_2 + zU_3) - y^2\,dz(xU_1 + yU_2 + zU_3)\\ &= x^2(x\,dx(U_1) + y\,dx(U_2) + z\,dx(U_3))\\ &\quad - y^2(x\,dz(U_1) + y\,dz(U_2) + z\,dz(U_3))\\ &\quad \text{(linearity of 1-forms)}\\ &= x^2(x + 0 + 0) - y^2(0 + 0 + z)\\ &= x^3 - y^2 z.\end{aligned}$$

(b) Similarly,

$$\begin{aligned}\phi(W) &= (x^2\,dx - y^2\,dz)((xy + yz)U_1 - yzU_2 - xyU_3)\\ &= x^2(xy + yz) - y^2(-xy)\\ &\quad \text{(method as first solution)}\\ &= xy(x^2 + xz - y^2).\end{aligned}$$

(c) There is no need to start from scratch with this one, we can use the two previous results and linearity:

$$\begin{aligned}\phi\left(\frac{1}{x}V + \frac{1}{y}W\right) &= \frac{1}{x}\phi(V) + \frac{1}{y}\phi(W) \quad \text{(linearity)}\\ &= \frac{1}{x}(x^3 - y^2 z) + \frac{1}{y}xy(x^2 + xz - y^2)\\ &= \frac{x^3 - y^2 z}{x} + x(x^2 + xz - y^2).\end{aligned}$$

Solution 5.3

These are all applications of the chain rule.

(a) Here,

$$d(f^5) = 5f^4\,df.$$

(b) This time,

$$\begin{aligned}d(\sqrt{f}) &= \frac{1}{2}f^{-1/2}\,df\\ &= \frac{df}{2\sqrt{f}}.\end{aligned}$$

(c) Finally,

$$\begin{aligned}d(\log(1+f^2)) &= \frac{1}{1+f^2}2f\,df\\ &= \frac{2f\,df}{1+f^2}.\end{aligned}$$

Solution 5.4

It is worth noting that the three partial derivatives of

$$(x^2 + y^2 + z^2)^{1/2}$$

will all have the same form because of the symmetry:

$$\begin{aligned}\frac{\partial(x^2+y^2+z^2)^{1/2}}{\partial x} &= \tfrac{1}{2}(x^2+y^2+z^2)^{-1/2}(2x)\\ &= \frac{x}{(x^2+y^2+z^2)^{1/2}}.\end{aligned}$$

Similarly,

$$\frac{\partial(x^2+y^2+z^2)^{1/2}}{\partial y} = \frac{y}{(x^2+y^2+z^2)^{1/2}},$$

$$\frac{\partial(x^2+y^2+z^2)^{1/2}}{\partial z} = \frac{z}{(x^2+y^2+z^2)^{1/2}}.$$

Hence

$$df = \frac{1}{(x^2+y^2+z^2)^{1/2}}(x\,dx + y\,dy + z\,dz).$$

Solution 5.5

In each case we check the two requirements: that ϕ be a real-valued function on tangent vectors and that it be linear at each point of $\mathbf{E}^3$.

In order to express the 1-forms in terms of dx, dy and dz, we use

$$dx(\mathbf{v}) = v_1, \quad dy(\mathbf{v}) = v_2, \quad dz(\mathbf{v}) = v_3$$

and the pointwise principle.

As usual, we use dx etc. and dx_1 etc. interchangeably.

(a) This is a 1-form. First, by the definition given, ϕ is a real-valued function on tangent vectors. Next,

$$\begin{aligned}\phi(a\mathbf{v}+b\mathbf{w}) &= (av_1+bw_1)-(av_3+bw_3)\\ &= a(v_1-v_3)+b(w_1-w_3)\\ &= a\phi(\mathbf{v})+b\phi(\mathbf{w}),\end{aligned}$$

so ϕ is linear at each point of $\mathbf{E}^3$.

We have

$$\phi(\mathbf{v}) = v_1 - v_3 = dx(\mathbf{v}) - dz(\mathbf{v}),$$

so, by the pointwise principle,

$$\phi = dx - dz.$$

(b) This is not a 1-form, although it is a real-valued function on tangent vectors. It fails to be linear, as the following counterexample shows. Let $\mathbf{v}$ be any tangent vector at the point $\mathbf{p} = (1,0,0)$. Then

$$\phi(\mathbf{v}) = 1 - 0 = 1.$$

However, $\phi(2\mathbf{v}) = 1 - 0 = 1$, since $2\mathbf{v}$ is based at the same point, thus

$$\phi(2\mathbf{v}) \neq 2\phi(\mathbf{v}).$$

(c) This is a 1-form. It is a real-valued function on tangent vectors and, at each point, p_1 and p_3 are fixed so

$$\begin{aligned}\phi(a\mathbf{v}+b\mathbf{w}) &= (av_1+bw_1)p_3+(av_2+bw_2)p_1\\ &= a(v_1p_3+v_2p_1)+b(w_1p_3+w_2p_1)\\ &= a\phi(\mathbf{v})+b\phi(\mathbf{w}),\end{aligned}$$

so ϕ is linear.

Here

$$\begin{aligned}\phi(\mathbf{v}) &= v_1p_3+v_2p_1\\ &= dx(\mathbf{v})\,z(\mathbf{p})+dy(\mathbf{v})\,x(\mathbf{p}),\end{aligned}$$

so

$$\phi = z\,dx + x\,dy.$$

(d) This is a 1-form. By the definition, $\mathbf{v}_{\mathbf{p}}[f]$ is a real number for any function f on $\mathbf{E}^3$, so ϕ is a real-valued function on tangent vectors. Also, Theorem 3.3(1) shows that

$$\begin{aligned}\phi(a\mathbf{v}+b\mathbf{w}) &= (a\mathbf{v}+b\mathbf{w})[x^2+y^2]\\ &= a\mathbf{v}[x^2+y^2]+b\mathbf{w}[x^2+y^2]\\ &= a\phi(\mathbf{v})+b\phi(\mathbf{w}),\end{aligned}$$

and ϕ is linear at each point.

We first evaluate $\phi(\mathbf{v})$ using

$$\mathbf{v}_{\mathbf{p}}[f] = \sum_{i=1}^{3}\frac{\partial f}{\partial x_i}(\mathbf{p})\,v_i$$

as

$$\begin{aligned}\phi(\mathbf{v}) &= \mathbf{v}_{\mathbf{p}}[x^2+y^2]\\ &= 2x(\mathbf{p})\,v_1+2y(\mathbf{p})\,v_2+0\,v_3\\ &= 2x(\mathbf{p})\,dx(\mathbf{v})+2y(\mathbf{p})\,dy(\mathbf{v}).\end{aligned}$$

Thus

$$\phi = 2x\,dx + 2y\,dy.$$

Note: You could simply have quoted *O'Neill*'s general result that df is a 1-form whose value on $\mathbf{v}_{\mathbf{p}}$ is $\mathbf{v}_{\mathbf{p}}[f]$. The last part follows from

$$df = \sum_{i=1}^{3}\frac{\partial f}{\partial x_i}\,dx_i.$$

(e) This is a 1-form. It is a real-valued function on tangent vectors. Also,

$$\begin{aligned}\phi(a\mathbf{v}+b\mathbf{w}) &= 0\\ &= 0+0\\ &= a\phi(\mathbf{v})+b\phi(\mathbf{w}).\end{aligned}$$

In this case it is clear that

$$\phi = 0\,dx + 0\,dy + 0\,dz.$$

(f) This is not a 1-form. It fails to be linear, and the same counterexample as used in part (b) will illustrate this failure. (Use any tangent vector based at $(1,0,0)$.)

Solution 6.1

In the first part we make use of the distributive properties of the wedge product, the alternation rule and the 'repeats are zero' property.

In the second part, we use the definition of exterior derivative.

(a) (i) We have

$$\begin{aligned}\phi\wedge\psi &= (yz\,dx+dz)\wedge(\sin z\,dx+\cos z\,dy)\\ &= yz\sin z\,dx\,dx+yz\cos z\,dx\,dy\\ &\quad+\sin z\,dz\,dx+\cos z\,dz\,dy\\ &= yz\cos z\,dx\,dy-\sin z\,dx\,dz-\cos z\,dy\,dz.\end{aligned}$$

(ii) Here

$$\begin{aligned}\psi\wedge\xi &= (\sin z\,dx+\cos z\,dy)\wedge(dy+z\,dz)\\ &= \sin z\,dx\,dy+z\sin z\,dx\,dz\\ &\quad+\cos z\,dy\,dy+z\cos z\,dy\,dz\\ &= \sin z\,dx\,dy+z\sin z\,dx\,dz+z\cos z\,dy\,dz.\end{aligned}$$

(iii) In this case

$$\begin{aligned}\xi\wedge\phi &= (dy+z\,dz)\wedge(yz\,dx+dz)\\ &= yz\,dy\,dx+dy\,dz+yz^2\,dz\,dx+z\,dz\,dz\\ &= -yz\,dx\,dy-yz^2\,dx\,dz+dy\,dz.\end{aligned}$$

(b) We use the definition, which amounts to saying 'to find the exterior derivative, find d of the coefficient functions'.

(i) We obtain

$$\begin{aligned}d\phi &= d(yz\,dx+dz)\\ &= d(yz)\,dx+d(1)\,dz\\ &= (0\,dx+z\,dy+y\,dz)dx+0\\ &= z\,dy\,dx+y\,dz\,dx\\ &= -z\,dx\,dy-y\,dx\,dz.\end{aligned}$$

(ii) Here

$$\begin{aligned} d\psi &= d(\sin z\, dx + \cos z\, dy) \\ &= d(\sin z)\, dx + d(\cos z)\, dy \\ &= \cos z\, dz\, dx - \sin z\, dz\, dy \\ &= -\cos z\, dx\, dz + \sin z\, dy\, dz. \end{aligned}$$

(iii) In this case

$$\begin{aligned} d\xi &= d(dy + z\, dz) \\ &= d(1)\, dy + dz\, dz \\ &= 0. \end{aligned}$$

Solution 6.2

We calculate the three terms individually. First,

$$\begin{aligned} d(\phi \wedge \psi) &= d\left(\frac{1}{y}\, dx \wedge z\, dy\right) \\ &= d\left(\frac{z}{y}\, dx\, dy\right) \\ &= d\left(\frac{z}{y}\right)\, dx\, dy \\ &= \left(\frac{-z}{y^2}\, dy + \frac{1}{y}\, dz\right)\, dx\, dy \\ &= \frac{-z}{y^2}\, dy\, dx\, dy + \frac{1}{y}\, dz\, dx\, dy \\ &= 0 - \frac{1}{y}\, dx\, dz\, dy \\ &= -\left(-\frac{1}{y}\, dx\, dy\, dz\right) \\ &= \frac{1}{y}\, dx\, dy\, dz. \end{aligned}$$

Next,

$$\begin{aligned} d\phi \wedge \psi &= d\left(\frac{1}{y}\, dx\right) \wedge z\, dy \\ &= -\frac{1}{y^2}\, dy\, dx \wedge z\, dy \\ &= 0 \quad \text{because } dy \text{ is repeated.} \end{aligned}$$

Finally,

$$\begin{aligned} -\phi \wedge d\psi &= -\frac{1}{y}\, dx \wedge d(z\, dy) \\ &= -\frac{1}{y}\, dx \wedge dz\, dy \\ &= -\frac{1}{y}\, dx\, dz\, dy \\ &= \frac{1}{y}\, dx\, dy\, dz. \end{aligned}$$

Comparing these results, we see that

$$d(\phi \wedge \psi) = d\phi \wedge \psi - \phi \wedge d\psi$$

in this case.

Solution 6.3

This solution hinges on the fact that 'mixed' second derivatives are independent of the order of differentiation (at least for the well-behaved functions considered in this course). That is,

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).$$

We know that

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz.$$

So, finding the differential of each coefficient,

$$\begin{aligned} d(df) &= \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial x}\, dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial x}\, dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial x}\, dz\right)\, dx \\ &\quad + \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\, dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial y}\, dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial y}\, dz\right)\, dy \\ &\quad + \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial z}\, dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial z}\, dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial z}\, dz\right)\, dz \\ &= \frac{\partial}{\partial y}\frac{\partial f}{\partial x}\, dy\, dx + \frac{\partial}{\partial x}\frac{\partial f}{\partial y}\, dx\, dy \\ &\quad + \frac{\partial}{\partial z}\frac{\partial f}{\partial x}\, dz\, dx + \frac{\partial}{\partial x}\frac{\partial f}{\partial z}\, dx\, dz \\ &\quad + \frac{\partial}{\partial z}\frac{\partial f}{\partial y}\, dz\, dy + \frac{\partial}{\partial y}\frac{\partial f}{\partial z}\, dy\, dz \\ &= 0 + 0 + 0 = 0. \end{aligned}$$

The final collapse is caused by the remark before the solution and the alternation rule.

Applying the above result gives

$$\begin{aligned} d(f\, dg) &= df \wedge dg + f\, d(dg) \quad \text{(Theorem 6.4(2))} \\ &= df \wedge dg + 0 \\ &= df \wedge dg. \end{aligned}$$

Solution 6.4

In this solution we apply the various parts of Theorem 6.4 together with linearity and the alternation rule. The result from the last solution is also useful.

(a) We obtain

$$\begin{aligned} d(f\, dg + g\, df) &= df \wedge dg + f\, d(dg) + dg \wedge df + g\, d(df) \\ &= df \wedge dg + 0 + dg \wedge df + 0 \\ &= df \wedge dg + dg \wedge df \\ &= 0. \end{aligned}$$

Note: If you happen to recognize that

$$f\, dg + g\, df = d(fg),$$

then the result can be obtained quickly:

$$d(f\, dg + g\, df) = d(d(fg)) = 0.$$

(b) In this case, we have

$$\begin{aligned} &d((f-g)(df + dg)) \\ &= d(f-g) \wedge (df + dg) + (f-g)d(df + dg) \\ &= (df - dg) \wedge (df + dg) + (f-g)(d(df) + d(dg)) \\ &= df \wedge df + df \wedge dg - dg \wedge df - dg \wedge dg + (f-g)(0+0) \\ &= df \wedge dg + df \wedge dg \\ &= 2\, df \wedge dg. \end{aligned}$$

(c) Here we apply Theorem 6.4(3), then Theorem 6.4(2):

$$\begin{aligned} d(f\, dg \wedge g\, df) &= d(f\, dg) \wedge (g\, df) - (f\, dg) \wedge d(g\, df) \\ &= (df \wedge dg + f\, d(dg)) \wedge (g\, df) - ((f\, dg) \wedge (dg \wedge df + g\, d(df))) \\ &= (df \wedge dg + 0) \wedge (g\, df) - (f\, dg) \wedge (dg \wedge df + 0) \\ &= 0. \end{aligned}$$

(Each term has a repeated differential.)

(d) Here we have

$$\begin{aligned} &d(gf\, df) + d(f\, dg) \\ &= d(gf) \wedge df + gf\, d(df) + df \wedge dg + f\, d(dg) \\ &= (f\, dg + g\, df) \wedge df + 0 + df \wedge dg + 0 \\ &= f\, dg \wedge df + 0 + df \wedge dg \\ &= (1-f)\, df \wedge dg. \end{aligned}$$

Solution 6.5

Following the hint, if we start with

$$\phi = f_1\,dx_1 + f_2\,dx_2 + f_3\,dx_3,$$

then

$$d\phi = df_1 \wedge dx_1 + df_2 \wedge dx_2 + df_3 \wedge dx_3,$$

by the definition of exterior product. Now, consider the exterior derivative of each of the terms on the right-hand side of the last equation. For the ith term,

$$\begin{aligned} d(df_i \wedge dx_i) &= d(df) \wedge dx_i - df \wedge d(dx_i) \\ &= 0 + 0 = 0. \end{aligned}$$

Both terms are zero because $d(dg) = 0$ for any function g and so the second derivatives of both f and x_i are zero. Hence $d(d\phi) = 0$.

Solution 6.6

Because the functions are $\mathbf{E}^2 \longrightarrow \mathbf{R}$, we have the differentials

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy,$$

$$dg = \frac{\partial g}{\partial x}\,dx + \frac{\partial g}{\partial y}\,dy.$$

Hence

$$\begin{aligned} df \wedge dg &= (\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy) \wedge (\frac{\partial g}{\partial x}\,dx + \frac{\partial g}{\partial y}\,dy) \\ &= 0 + \frac{\partial f}{\partial x}\frac{\partial g}{\partial y}\,dx\,dy + \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\,dy\,dx + 0 \\ &= (\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x})\,dx\,dy \\ &= \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{vmatrix}\,dx\,dy. \end{aligned}$$

Note: This result does imply the alternation rule because $dg \wedge df$ will give rise to a determinant consisting of the same entries as above but with the columns interchanged. As you can easily verify, interchanging two columns of a determinant changes the sign. Hence

$$dg \wedge df = -df \wedge dg.$$

Solution 7.1

We have used x, y, z for the coordinate functions. Answers expressed in terms of x_1 etc. are equally acceptable.

(a) Since

$$-3\mathbf{p} = (-3p_1, -3p_2, -3p_3),$$

we have

$$F = (-3x, -3y, -3z).$$

(b) Here we have

$$F = (e^{xy}, z + 2y, x^2).$$

Solution 7.2

(a) If $\mathbf{p}$ lies on the line $x = 1$, then $\mathbf{p} = (1, p_2)$. Thus

$$F(\mathbf{p}) = (2p_2, 1 - p_2^2).$$

There are various ways of deciding what curve in $\mathbf{E}^2$ is represented by

$$p_2 \longmapsto (2p_2, 1 - p_2^2).$$

One way is to see that the coordinates are related by

$$y = 1 - \left(\frac{x}{2}\right)^2,$$

and so the (route of the) curve is a parabola.

(b) A similar approach shows that if $\mathbf{p} = (p_1, -1)$, then

$$F(\mathbf{p}) = (-2p_1, p_1^2 - 1),$$

so the curve has equation

$$y = \left(\frac{-x}{2}\right)^2 - 1 = \left(\frac{x}{2}\right)^2 - 1.$$

This is also a parabola.

Solution 7.3

(a) Suppose that $\mathbf{v_p}$ is a tangent vector to $\mathbf{E}^3$. Then, by the definition,

$$F_*(\mathbf{v_p}) = (F(\mathbf{p} + t\mathbf{v}))'(0)_{F(\mathbf{p})}.$$

Now,

$$\begin{aligned} F(\mathbf{p} + t\mathbf{v}) &= F(p_1 + tv_1, p_2 + tv_2, p_3 + tv_3) \\ &= ((p_1 + tv_1) - (p_2 + tv_2), (p_1 + tv_1) + (p_2 + tv_2), 2(p_3 + tv_3)) \\ &= ((p_1 - p_2) + t(v_1 - v_2), (p_1 + p_2) + t(v_1 + v_2), 2p_3 + 2tv_3), \end{aligned}$$

$$(F(\mathbf{p} + t\mathbf{v}))' = (v_1 - v_2, v_1 + v_2, 2v_3),$$

$$(F(\mathbf{p} + t\mathbf{v}))'(0) = (v_1 - v_2, v_1 + v_2, 2v_3),$$

$$F(\mathbf{p}) = (p_1 - p_2, p_1 + p_2, 2p_3).$$

Hence

$$F_*(\mathbf{v_p}) = (v_1 - v_2, v_1 + v_2, 2v_3)_{(p_1 - p_2, p_1 + p_2, 2p_3)}.$$

Note that *in this case* we also have

$$F_*(\mathbf{v_p}) = F(\mathbf{v})_{F(\mathbf{p})}.$$

(b) Using the standard notation for components of $\mathbf{p}$ and $\mathbf{q}$, we have

$$\begin{aligned} F(a\mathbf{p} + b\mathbf{q}) &= F(ap_1 + bq_1, ap_2 + bq_2, ap_3 + bq_3) \\ &= ((ap_1 + bq_1) - (ap_2 + bq_2), (ap_1 + bq_1) + (ap_2 + bq_2), \\ &\qquad 2(ap_3 + bq_3)) \\ &= (a(p_1 - p_2) + b(q_1 - q_2), a(p_1 + p_2) + b(q_1 + q_2), a2p_3 + b2q_3) \\ &= a(p_1 - p_2, p_1 + p_2, 2p_3) + b(q_1 - q_2, q_1 + q_2, 2q_3) \\ &= aF(\mathbf{p}) + bF(\mathbf{q}). \end{aligned}$$

(c) We apply the definition of G_* and the information that G is linear:

$$\begin{aligned} G_*(\mathbf{v_p}) &= (G(\mathbf{p} + t\mathbf{v}))'(0)_{G(\mathbf{p})} \\ &= (G(\mathbf{p}) + tG(\mathbf{v}))'(0)_{G(\mathbf{p})} \quad \text{(linearity)} \\ &= (0 + G(\mathbf{v}))(0)_{G(\mathbf{p})} \quad \text{(differentiating w.r.t. } t\text{)} \\ &= G(\mathbf{v})_{G(\mathbf{p})}. \end{aligned}$$

Solution 7.4

(a) Since

$$\frac{\partial(x + y)}{\partial x} = 1$$

etc., we have that the Jacobian for F is

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(b) Similarly, the Jacobian for G is

$$\begin{pmatrix} 2x & 2y & 0 \\ 0 & 2y & 2z \end{pmatrix}.$$

Evaluating this at F gives

$$\begin{pmatrix} 2(x + y) & 2(x - 2y) & 0 \\ 0 & 2(x - 2y) & 2(x + y + z) \end{pmatrix}$$

since $x(F) = x(x+y, x-2y, x+y+z) = x+y$, etc.

(c) Combining the last two parts gives the Jacobian for $G \circ F$ as

$$\begin{pmatrix} 2(x+y) & 2(x-2y) & 0 \\ 0 & 2(x-2y) & 2(x+y+z) \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2(2x-y) & 2(-x+5y) & 0 \\ 2(2x-y+z) & 2(-x+5y+z) & 2(x+y+z) \end{pmatrix}.$$

ISBN 0 7492 4770 3